Eastern Economy Edition

ENGINEERING MATHEMATICS

Volume II

A Course in
Linear Algebra and Its Applications

J.P. Sharma

Mahesh A. Yeolekar

PHI

Engineering Mathematics

Volume II

[As Per Gujarat Technological University (GTU) Syllabus]

J.P. Sharma
Principal
Hasmukh Goswami College of Engineering (HGCE)
Ahmedabad

Mahesh A. Yeolekar
Head, Department of Mathematics
Hasmukh Goswami College of Engineering (HGCE)
Ahmedabad

PHI Learning Private Limited
New Delhi-110001
2011

ENGINEERING MATHEMATICS, Volume II
J.P. Sharma and Mahesh A. Yeolekar

ISBN-978-81-203-4248-4

Published by Asoke K. Ghosh, PHI Learning Private Limited, M-97, Connaught Circus, New Delhi-110001 and Printed by Baba Barkha Nath Printers, Bahadurgarh, Haryana-124507.

Contents

Preface

This Volume II of Engineering Mathematics is a companion volume to our Volume I of Engineering Mathematics published in 2009 for the first-year, first-semester B.Tech/B.E. students of all disciplines of engineering. The Volume II in continuation, is designed for the first-year, second-semester B.Tech/B.E. students. Whereas the Volume I covered the topics in calculus, Volume II presents the mathematical concepts, tools and techniques of **linear algebra**. The knowledge of linear algebra is of vital importance to engineers as it finds widespread applications in the solution of problems encountered in engineering profession.

This book, therefore, introduces students to the related concepts, rules and use of **matrices** and **vector spaces** in the solution of linear systems of equations which appear frequently as models of various engineering problems. Separate chapters are devoted to a thorough study of **linear transformations, linear product spaces** and **eigenvalue problems** in connection with matrices. The book makes liberal use of solved examples and provides plenty of exercises for homework in order that the students can apply these mathematical methods to the successful solution of real problems.

The book fully conforms to the syllabus of Gujarat Technological University. Being a basic book in linear algebra, it will be useful to the students of engineering in all other universities as well.

We express our sincere thanks to Prof. H.P. Singh, DDIT, Nadiad and all our colleagues at Hasmukh Goswami College of Engineering, Ahmedabad and Atmiya Institute of Technology and Science, Rajkot for inspiring and motivating us to write this book.

We are indebted to all those who contributed directly or indirectly to the completion of this book; in particular, we would like to mention Shri Hasmukh Goswami (Turstee, Monark Education Trust), Shri Janakbhai Khandwala (Executive Director, HGCE), and Dr. M.B. Patel (Principal, HGCE) for their moral support in completion of this project.

We wish to thank the editorial and production team of PHI Learning for their careful processing of the manuscript.

Suggestions for the improvement of this book will be gratefully acknowledged.

J.P. Sharma
Mahesh A. Yeolekar

1 Matrices and Linear Equations

1.1 THE THEORY OF MATRICES

The theory of matrices was developed as a means to solve simultaneous, linear algebraic equations. Presently, it spans the entire spectrum of physical sciences. In this chapter, we will define the notation, the terminology and central ideas of the theory of matrices.

Definition: Matrix

A rectangular array of numbers of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is called an $m \times n$ matrix, with m rows and n columns. Here $m \times n$ (read "m by n") is called the *order* of the matrix. We count rows from the top and columns from the left. Hence

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{bmatrix}$$

represent respectively the ith row and the jth column of the above matrix, and a_{ij} represents the entry in the matrix on the ith row and jth column.

Let us consider a 4×3 matrix as given below. It is the practice to denote matrices by single capital bold face letters such as **A**, **B**, **C** ... .

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 5 & 4 \\ 3 & 9 & 10 \\ 2 & -3 & 7 \end{bmatrix}$$

Here $\begin{bmatrix} 0 & 5 & 4 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 4 \\ 10 \\ 7 \end{bmatrix}$

represent respectively the 2nd row and the 3rd column of the matrix, and 4 represents the entry in the matrix on the 2nd row and the 3rd column.

Matrix Operations

We now consider arithmetic operations involving matrices. First of all, let us study the problem of addition of matrices.

Definition: Addition

Suppose that the two matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ . & . & \cdots & . \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

both have m rows and n columns, that is, both **A** and **B** are of the same size or order. Then, we can add them as follows

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ . & . & \cdots & . \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

and call this result the sum of the two matrices **A** and **B**.

EXAMPLE 1.1 Find the sum of matrices $\mathbf{A} = \begin{bmatrix} 1 & -2 & 6 \\ 4 & 8 & 3 \\ -1 & 9 & 0 \\ 7 & 3 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 4 & -5 \\ 6 & 1 & 3 \\ 9 & 7 & 9 \\ 2 & -1 & 5 \end{bmatrix}$.

Solution: By the definition of addition of matrices,

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} 1+0 & -2+4 & 6-5 \\ 4+6 & 8+1 & 3+3 \\ -1+9 & 9+7 & 0+9 \\ 7+2 & 3-1 & 1+5 \end{bmatrix}=\begin{bmatrix} 1 & 2 & 1 \\ 10 & 9 & 6 \\ 8 & 16 & 9 \\ 9 & 2 & 6 \end{bmatrix}$$

EXAMPLE 1.2 Find the sum of matrices $\mathbf{A}=\begin{bmatrix} 1 & -1 \\ 0 & 4 \\ 11 & 3 \end{bmatrix}$ and $\mathbf{B}=\begin{bmatrix} 0 & 5 \\ -6 & 2 \end{bmatrix}$ if it exists.

Solution: By the definition, the addition operation exists only when both matrices have the same numbers of rows and columns, that is, both matrices have the same order. But in the case of the given matrices of this example, the order of **A** is 3×2 and the order of **B** is 2×2, so the addition between **A** and **B** does not exist.

Theorem 1.1 [Properties of Addition]

Suppose that **A**, **B**, **C** are all $m \times n$ matrices. Further suppose that **0** represents the $m \times n$ matrix with all entries zero. Then, we have the following properties for matrix addition.

(i) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$
(ii) $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$
(iii) $\mathbf{A}+\mathbf{0}=\mathbf{A}$
(iv) There is an $m \times n$ matrix $\mathbf{A}'$ such that $\mathbf{A}+\mathbf{A}'=\mathbf{0}$

Definition: Scalar Multiplication

Suppose that the matrix

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

has m rows and n columns, and that c is any scalar (that is, any number). Then, we can write

$$c\mathbf{A}=\begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ . & . & \cdots & . \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

and call this result the product of the matrix **A** by the scalar c.

EXAMPLE 1.3 Find $2\mathbf{A}$, $3\mathbf{A}$, $-\mathbf{A}$, $3\mathbf{A}-\mathbf{A}$ for the matrix $\mathbf{A}=\begin{bmatrix} 1 & 4 & 5 \\ -5 & 3 & 1 \\ 0 & 9 & 7 \end{bmatrix}$.

Solution: By the definition of scalar multiplication,

$$3\mathbf{A} = \begin{bmatrix} 3 & 12 & 15 \\ -15 & 9 & 3 \\ 0 & 27 & 21 \end{bmatrix}$$

$$-\mathbf{A} = \begin{bmatrix} -1 & -4 & -5 \\ 5 & -3 & -1 \\ 0 & -9 & -7 \end{bmatrix}$$

$$3\mathbf{A} - \mathbf{A} = \begin{bmatrix} 3-1 & 12-4 & 15-5 \\ -15+5 & 9-3 & 3-1 \\ 0+0 & 27-9 & 21-7 \end{bmatrix} = \begin{bmatrix} 2 & 8 & 10 \\ -10 & 6 & 2 \\ 0 & 18 & 14 \end{bmatrix}$$

$$2\mathrm{A} = \begin{bmatrix} 2 & 8 & 10 \\ -10 & 6 & 2 \\ 0 & 18 & 14 \end{bmatrix}$$

Theorem 1.2 [Properties of Scalar Multiplication]

Suppose that **A**, **B** are both $m \times n$ matrices, and that c, d are any two scalars. Further suppose that **0** represents the $m \times n$ matrix with all entries zero. Then, we have the following properties for scalar multiplication.

(i) $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$
(ii) $(c + d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}$
(iii) $0\mathbf{A} = \mathbf{0}$
(iv) $c(d\mathbf{A}) = (cd)\mathbf{A}$

Definition: Matrix Multiplication

Suppose that

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ . & . & \cdots & . \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

are respectively an $m \times n$ matrix and an $n \times p$ matrix. Then the matrix product (matrix multiplication) $\mathbf{AB} = \mathbf{C}$ is given by the $m \times p$ matrix **C**. That is,

$$\mathbf{AB} = \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ . & . & \cdots & . \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

where for every $i = 1, \ldots, m$ and $j = 1, \ldots, p$, we have

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$$

Remark: Note that the matrix multiplication of two matrices exists only when the number of columns of the first matrix must be equal to the number of rows of the second matrix.

EXAMPLE 1.4 Find **AB**, **BA** if they exist for the matrices $\mathbf{A} = \begin{bmatrix} 1 & 4 & -2 \\ 0 & 1 & 3 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 1 & -1 & 2 \\ 1 & 5 & 2 & 7 \\ 3 & -6 & 3 & 4 \end{bmatrix}$.

Solution: Note that **A** is a 2×3 matrix and **B** is a 3×4 matrix, so that the product **AB** is a 2×4 matrix. Let us calculate the product.

$$\mathbf{AB} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \end{bmatrix}$$

Consider first of all c_{11}. To calculate this entry, we need only the 1st row of **A** and the 1st column of **B**. So, let us hide all the unnecessary information and write

$$\begin{bmatrix} 1 & 4 & -2 \\ \times & \times & \times \end{bmatrix} \begin{bmatrix} 0 & \times & \times & \times \\ 1 & \times & \times & \times \\ 3 & \times & \times & \times \end{bmatrix} = \begin{bmatrix} c_{11} & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}$$

From the definition, we have

$$c_{11} = 1 \times 0 + 4 \times 1 + (-2) \times 3 = -2$$

Consider next the entry c_{12}. To calculate this, we need only the 1st row of **A** and the 2nd column of **B**. So, let us hide all the unnecessary information, and write

$$\begin{bmatrix} 1 & 4 & -2 \\ \times & \times & \times \end{bmatrix} \begin{bmatrix} \times & 1 & \times & \times \\ \times & 5 & \times & \times \\ \times & -6 & \times & \times \end{bmatrix} = \begin{bmatrix} \times & c_{12} & \times & \times \\ \times & \times & \times & \times \end{bmatrix}$$

From the definition, we have

$$c_{12} = 1 \times 1 + 4 \times 5 + (-2) \times (-6) = 33$$

Now consider c_{23}. To calculate this entry, we need only the 2nd row of **A** and the 3rd column of **B**. So, let us hide all the unnecessary information, and write

$$\begin{bmatrix} \times & \times & \times \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} \times & \times & -1 & \times \\ \times & \times & 2 & \times \\ \times & \times & 3 & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & c_{23} & \times \end{bmatrix}$$

From the definition, we have

$$c_{23} = 0 \times (-1) + 1 \times 2 + (3) \times (3) = 11$$

Similarly, we can calculate the other entries of the product matrix **AB** to get

$$\mathbf{AB} = \begin{bmatrix} -2 & 33 & 1 & 22 \\ 10 & -13 & 11 & 19 \end{bmatrix}$$

The product **BA** does not exist because the number of columns of **B** is not equal to the number of rows of **A**.

Remark: Note from Example 1.4 that matrix multiplication is not a commutative operation, that is, $\mathbf{AB} \neq \mathbf{BA}$.

Theorem 1.3 [Associative Laws]

Suppose that $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{B}$ is an $n \times p$ matrix and $\mathbf{C}$ is a $p \times r$ matrix. Then

(i) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
(ii) $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B} = \mathbf{A}(c\mathbf{B})$, where c is any scalar.

Theorem 1.4 [Distributive Laws]

(i) Suppose that $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ and $\mathbf{C}$ are $n \times p$ matrices. Then $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.
(ii) Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices and $\mathbf{C}$ is an $n \times p$ matrix. Then $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$.

Transpose of a Matrix

Definition: Transposition

Let $\mathbf{A}$ be an $m \times n$ matrix. Then $\mathbf{A}^{\mathrm{T}}$, the transpose of $\mathbf{A}$, is the matrix obtained by interchanging the rows and columns of $\mathbf{A}$. In other words if $\mathbf{A} = [a_{ij}]$, then $\mathbf{A}^{\mathrm{T}} = [a_{ji}]$. Consequently $\mathbf{A}^{\mathrm{T}}$ is an $n \times m$ matrix.

EXAMPLE 1.5 Find the transpose of a matrix $\mathbf{A} = \begin{bmatrix} 2 & 8 & 10 \\ -10 & 6 & 2 \\ 0 & 18 & 14 \end{bmatrix}$.

Solution: By the definition of transpose, $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 2 & -10 & 0 \\ 8 & 6 & 18 \\ 10 & 2 & 14 \end{bmatrix}$.

Theorem 1.5 [Properties of Transpose]

(i) $(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{A}$
(ii) $(\mathbf{A} \pm \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \pm \mathbf{B}^{\mathrm{T}}$ if $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices
(iii) $(c\mathbf{A})^{\mathrm{T}} = c\mathbf{A}^{\mathrm{T}}$ where c is a scalar
(iv) $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ if $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times p$.

Trace of a Matrix

Definition: Trace

Let $\mathbf{A} = [a_{ij}]$ be an $n \times n$ matrix, that is, a square matrix. Then the *trace* of $\mathbf{A}$, written as tr $\mathbf{A}$, is the sum of the diagonal elements of $\mathbf{A}$, that is,

$$\operatorname{tr} \mathbf{A} = \sum_{k=1}^{n} a_{kk}$$

Theorem 1.6 [Properties of Trace]

Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ matrices. Then

(i) $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B}$
(ii) $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$

Remark: We cannot find the trace of a non-square matrix.

EXAMPLE 1.6 Find the trace of the following matrices $\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 \\ 3 & 0 & 1 \\ 6 & 5 & 9 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 4 & 1 \\ -3 & -5 & 5 \\ 7 & 2 & -2 \end{bmatrix}$.

Also verify the results of Theorem 1.6.

Solution: By the definition of trace,

$$\text{tr}\,\mathbf{A} = a_{11} + a_{22} + a_{33} = 1 + 0 + 9 = 10$$

$$\text{tr}\,\mathbf{B} = b_{11} + b_{22} + b_{33} = 0 - 5 - 2 = -7$$

$$\text{tr}\,\mathbf{A} + \text{tr}\,\mathbf{B} = 10 - 7 = 3$$

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -5 & 6 \\ 13 & 7 & 7 \end{bmatrix}$$

$$\text{tr}\,(\mathbf{A} + \mathbf{B}) = 1 - 5 + 7 = 3$$

$\therefore$
$$\text{tr}\,(\mathbf{A} + \mathbf{B}) = \text{tr}\,\mathbf{A} + \text{tr}\,\mathbf{B} \qquad \text{(i)}$$

Now,

$$\mathbf{AB} = \begin{bmatrix} 6 & 14 & -9 \\ 7 & 14 & 1 \\ 48 & 17 & 13 \end{bmatrix}; \qquad \mathbf{BA} = \begin{bmatrix} 18 & 5 & 13 \\ 12 & 31 & 40 \\ 1 & -24 & -16 \end{bmatrix}$$

$$\text{tr}\,(\mathbf{AB}) = 6 + 14 + 13 = 33; \qquad \text{tr}\,(\mathbf{BA}) = 18 + 31 - 16 = 33$$

Hence
$$\text{tr}\,(\mathbf{AB}) = \text{tr}\,(\mathbf{BA}).$$

Elementary Row Operations for Matrices

Definition: ***Elementary Row Operations***

There are three types of elementary row operations that can be performed on matrices:

(i) Interchanging two rows:

$R_i \leftrightarrow R_j$ interchanges rows i and j.

(ii) Multiplying a row by a nonzero scalar:

$R_i \rightarrow tR_i$ multiplies row i by the nonzero scalar t.

(iii) Adding a multiple of one row to another row:

$R_j \rightarrow R_j + tR_i$ adds t times row i to row j.

Definition: ***Row Equivalence***

Matrix **A** is row-equivalent to matrix **B** if **B** is obtained from **A** by a sequence of elementary row operations.

EXAMPLE 1.7 Are $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & 4 & 0 \\ 1 & -1 & 2 \\ 4 & -1 & 5 \end{bmatrix}$ equivalent matrices?

Solution:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$R_2 \to R_2 + 2R_3 \quad \sim \quad \begin{bmatrix} 1 & 2 & 0 \\ 4 & -1 & 5 \\ 1 & -1 & 2 \end{bmatrix}$$

$$R_2 \leftrightarrow R_3 \quad \sim \quad \begin{bmatrix} 1 & 2 & 0 \\ 1 & -1 & 2 \\ 4 & -1 & 5 \end{bmatrix}$$

$$R_1 \to 2R_1 \quad \sim \quad \begin{bmatrix} 2 & 4 & 0 \\ 1 & -1 & 2 \\ 4 & -1 & 5 \end{bmatrix} = \mathbf{B}$$

Thus **A** is row-equivalent to **B**. Clearly **B** is also row-equivalent to **A**, which can be proved by performing the inverse row-operations $R_1 \to \frac{1}{2} R_1$, $R_2 \leftrightarrow R_3$, $R_2 \to R_2 - 2R_3$ on **B**. Thus, **A** and **B** are equivalent matrices.

Definition: Row-Echelon Form

A matrix is in row-echelon form if it satisfies the following properties:

(i) All the zero rows (if any) are at the bottom of the matrix.
(ii) If the first nonzero entry in a nonzero row is a 1, it is called the *leading* 1 of a nonzero row.
(iii) If two successive rows are nonzero, the second row starts with more zeros than the first one (moving from left to right).

For example, here are some matrices

$$\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 & * \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 1 & * \end{bmatrix}, \begin{bmatrix} 0 & 1 & * \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

that are in row-echelon form.

The following matrices are not in row-echelon form.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

Remark: The zero matrix of any size is always in row-echelon form.

EXAMPLE 1.8 Find the row-echelon form of the matrix $\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ -2 & 3 & 0 \end{bmatrix}$.

Solution: $\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ -2 & 3 & 0 \end{bmatrix}$

$$R_3 \to R_3 + R_1 \quad \sim \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ 0 & 7 & 6 \end{bmatrix}$$

$$R_1 \to \frac{1}{2} R_1 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 9 \\ 0 & 7 & 6 \end{bmatrix} \qquad R_2 \to R_2 - R_1 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & 7 & 6 \end{bmatrix}$$

$$R_2 \leftrightarrow R_3 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 7 & 6 \\ 0 & 0 & 6 \end{bmatrix} \qquad \begin{matrix} R_2 \to \dfrac{1}{7} R_2 \\ R_3 \to \dfrac{1}{6} R_3 \end{matrix} \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & \dfrac{6}{7} \\ 0 & 0 & 1 \end{bmatrix}$$

It is a row-echelon form of the given matrix **A**.

Definition: Reduced Row-Echelon Form

A matrix is in reduced row-echelon form if it satisfies the following properties:

(i) It is in the row-echelon form.
(ii) The leading (leftmost nonzero) entry in each nonzero row is 1.
(iii) All other elements of the column in which the leading entry 1 occurs are zeros.

Remark: Every matrix has a unique reduced row-echelon form but it may have different row-echelon forms.

For example, here are some matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & * \\ 0 & 1 & * \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & * \\ 0 & 1 & * \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

that are in reduced row-echelon form.

The following matrices are not in reduced row-echelon form.

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

Remark: The zero matrix of any size is always in reduced row-echelon form.

EXAMPLE 1.9 Find the reduced row-echelon form of the matrix $\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ -2 & 3 & 0 \end{bmatrix}$.

Solution: $\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ -2 & 3 & 0 \end{bmatrix}$ $\quad R_3 \to R_3 + R_1 \quad \sim \begin{bmatrix} 2 & 4 & 6 \\ 1 & 2 & 9 \\ 0 & 7 & 6 \end{bmatrix}$

$$R_1 \to \frac{1}{2} R_1 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 9 \\ 0 & 7 & 6 \end{bmatrix} \qquad R_2 \to R_2 - R_1 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & 7 & 6 \end{bmatrix}$$

$$R_2 \leftrightarrow R_3 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 7 & 6 \\ 0 & 0 & 6 \end{bmatrix} \qquad R_2 \to R_2 - R_3 \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 7 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\begin{matrix} R_2 \to \frac{1}{7} R_2 \\ R_3 \to \frac{1}{6} R_3 \end{matrix} \quad \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_1 \to R_1 - 2R_2 \quad \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$R_1 \to R_1 - 3R_3 \quad \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. It is the reduced row-echelon form of the given matrix **A**.

EXAMPLE 1.10 Find the reduced row-echelon form of the matrix $\mathbf{B} = \begin{bmatrix} 1 & -5 & 7 \\ 6 & 9 & 7 \\ -9 & 6 & 1 \end{bmatrix}$.

Solution: $\mathbf{B} = \begin{bmatrix} 1 & -5 & 7 \\ 6 & 9 & 7 \\ -9 & 6 & 1 \end{bmatrix}$ $\quad R_2 \to R_2 - 6R_1 \quad \sim \begin{bmatrix} 1 & -5 & 7 \\ 0 & 39 & -35 \\ -9 & 6 & 1 \end{bmatrix}$

$$R_3 \to R_3 + 9R_1 \quad \sim \begin{bmatrix} 1 & -5 & 7 \\ 0 & 39 & -35 \\ 0 & 39 & 64 \end{bmatrix} \qquad R_3 \to R_3 - R_2 \quad \sim \begin{bmatrix} 1 & -5 & 7 \\ 0 & 39 & -35 \\ 0 & 0 & 99 \end{bmatrix}$$

$$R_3 \to \frac{1}{99} R_3 \quad \sim \begin{bmatrix} 1 & -5 & 7 \\ 0 & 39 & -35 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{matrix} R_2 \to R_2 + 35R_3 \\ R_1 \to R_1 - 7R_3 \end{matrix} \quad \sim \begin{bmatrix} 1 & -5 & 0 \\ 0 & 39 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$R_2 \to \frac{1}{39} R_2 \quad \sim \quad \begin{bmatrix} 1 & -5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_1 \to R_1 + 5R_2 \quad \sim \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. It is the reduced row-echelon form of the given matrix **B**.

Theorem 1.7 Row-echelon Form

If **R** is the reduced row-echelon form of an $n \times n$ matrix **A**, then either **R** is a row of zeros or **R** is the identity matrix $\mathbf{I}_n$.

EXERCISE SET 1

1. If $\mathbf{A} = \begin{bmatrix} 4 & 2 & 5 \\ 1 & 3 & -6 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \end{bmatrix}$, compute **A** + **B**, **B** + **A**, (– 4)**B**, 6**A**, 5**A** – 3**A**, and 2**B** – 3**B**.
2. If $\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix}$, find **AB** and **BA** (if it exists) and show that **AB** = **BA**.
3. If $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} -3 & 1 \\ 2 & 0 \end{bmatrix}$, verify (i) (**AB**)**C** = **A**(**BC**), (ii) **A**(**B** + **C**) =**AB** + **AC**, and (iii) (**A** + **B**)**C** = **AC** + **BC**.
4. If $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, calculate (i) $\mathbf{A}^2\mathbf{B} + \mathbf{B}^2\mathbf{A}$, (ii) prove that tr (**AB**) = tr (**BA**), and also (iii) prove that tr (**A** + **B**) = tr **A** + tr **B**.
5. $\begin{bmatrix} x+3 & 2y+x \\ z-1 & 4a-6 \end{bmatrix} = \begin{bmatrix} 0 & -7 \\ 3 & 2a \end{bmatrix}$, find the values of x, y, z and a which satisfy the matrix equation.
6. If $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & -2 \\ -2 & 0 \\ 2 & -1 \end{bmatrix}$, obtain the product **AB** and explain why **BA** is not defined and also calculate tr **A**.
7. Matrices **A** and **B** are such that $3\mathbf{A} - 2\mathbf{B} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$ and $-4\mathbf{A} + \mathbf{B} = \begin{bmatrix} -1 & 2 \\ -4 & 3 \end{bmatrix}$, find **A** and **B**.
8. If $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix}$, $\mathbf{A} - \mathbf{B} = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}$, calculate **AB**.
9. If $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix}$, show that (2**I** – **A**) (10**I** – **A**) = 9**I**.

10. If $\mathbf{A} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \cos^2\varphi & \cos\varphi\sin\varphi \\ \cos\varphi\sin\varphi & \sin^2\varphi \end{bmatrix}$, show that **AB** is zero where θ and φ differ by an odd multiply of $\dfrac{\pi}{2}$.

(**Hint:** $\theta - \varphi = (2n+1)\pi/2$

11. Determine which of the following matrices are in row-echelon form.

(i) $\begin{bmatrix} 1 & 4 & 2 & 5 \\ 0 & 1 & -5 & 3 \\ 0 & 0 & 2 & 4 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (v) $\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

12. Determine which of the following matrices are in reduced row-echelon form.

(i) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{bmatrix}$ (v) $\begin{bmatrix} 1 & -7 & 2 & 2 \\ 0 & 1 & 3 & 2 \end{bmatrix}$

13. Find the row-echelon form of the following matrices:

(i) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 2 \\ 2 & 6 & 5 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & 3 & 4 & -1 \\ 5 & 2 & 0 & -1 \\ -4 & 5 & 12 & -1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 6 & 8 & 7 & 5 \end{bmatrix}$

14. Find the reduced row-echelon form of the following matrices:

(i) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 0 & 5 & -10 \end{bmatrix}$ (iii) $\begin{bmatrix} 2 & 3 & -1 & -1 \\ 1 & -1 & -2 & -4 \\ 3 & 1 & 3 & -2 \\ 6 & 3 & 0 & -7 \end{bmatrix}$

1.2 SPECIAL MATRICES

In the preceding section, we defined the matrix and matrix operations. Now, in this section we will define some special types of matrices which will be used to develop other concepts of the matrix and the system of linear equations.

Row Matrix: A matrix with only one row is called a *row matrix*, that is, a matrix of order $1 \times n$, is called a *row matrix*. For example, the following row matrices are of the order 1×2, 1×3 and 1×5 respectively.

$$[1 \quad 2]; \quad [0 \quad 1 \quad -1]; \quad [4 \quad 0 \quad -6 \quad 2 \quad 1].$$

Column Matrix: A matrix with only one column is called a *column matrix*, that is, a matrix of order $m \times 1$, is called a *column matrix*. For example, the following column matrices are of the order 1×1, 3×1, and 4×1 respectively.

$$\begin{bmatrix} 1 \\ 9 \end{bmatrix}; \quad \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}; \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Null Matrix: An $m \times n$ matrix whose all entries are zeroes is called a *null matrix* or *zero matrix*. It is denoted by **0**. For example, the following null matrices are of the order 2×2, 3×3, and 4×3 respectively.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Square Matrix: An $m \times n$ matrix is called a *square matrix* if $m = n$. It is known as matrix of order n. For example, the following square matrices are of the order of 2, 3, and 4 respectively.

$$\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}; \quad \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \\ 6 & -8 & 2 \end{bmatrix}; \quad \begin{bmatrix} 5 & 6 & 1 & 9 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & -1 & -6 \\ 1 & 5 & 7 & 0 \end{bmatrix}.$$

Diagonal Matrix: A square matrix $\mathbf{A} = [a_{ij}]$ is called a *diagonal matrix* if $a_{ij} = 0$ when $i \neq j$. The entries a_{ii} are known as the diagonal entries of **A**. For example, the diagonal entries of the following square matrices are (1, 4), (–6, 2, 4) and (5, 1, 4, 2) respectively.

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}; \quad \begin{bmatrix} -6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}; \quad \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Identity Matrix: An $n \times n$ diagonal matrix is called an *identity matrix* if it has all 1s on the main diagonal. It is denoted by **I**.

In other words, $\mathbf{I} = [a_{ij}]$ is identity matrix if

$$a_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

For example, the following matrices are identity matrices of order 2×2, 3×3 and 4×4 respectively.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Remark: For every $n \times n$ matrix $\mathbf{A}$, we have $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.

Symmetric Matrices: A square matrix $\mathbf{A} = [a_{ij}]$ is *symmetric* if $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, that is if $[a_{ij}] = [a_{ji}]$. For example,

(i) If $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$, then $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix} = \mathbf{A}$.

(ii) If $\mathbf{B} = \begin{bmatrix} 4 & -5 & 0 \\ -5 & 1 & 3 \\ 0 & 3 & 1 \end{bmatrix}$, then $\mathbf{B}^{\mathrm{T}} = \begin{bmatrix} 4 & -5 & 0 \\ -5 & 1 & 3 \\ 0 & 3 & 1 \end{bmatrix} = \mathbf{B}$.

(iii) If $\mathbf{C} = \begin{bmatrix} 0 & 2 & 3 & 8 \\ 2 & -8 & 0 & 7 \\ 3 & 0 & -5 & 6 \\ 8 & 7 & 6 & 1 \end{bmatrix}$, then $\mathbf{C}^{\mathrm{T}} = \begin{bmatrix} 0 & 2 & 3 & 8 \\ 2 & -8 & 0 & 7 \\ 3 & 0 & -5 & 6 \\ 8 & 7 & 6 & 1 \end{bmatrix} = \mathbf{C}$.

Skew-symmetric Matrices: A square matrix $\mathbf{A} = [a_{ij}]$ is *skew symmetric* if $\mathbf{A}^{\mathrm{T}} = -\mathbf{A}$, that is

$$a_{ij} = \begin{cases} -a_{ji}, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases}$$

For example:

(i) If $\mathbf{A} = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$, then $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 0 & -3 \\ 3 & 0 \end{bmatrix} = -\mathbf{A}$

(ii) If $\mathbf{B} = \begin{bmatrix} 0 & 5 & 0 \\ -5 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix}$, then $\mathbf{B}^{\mathrm{T}} = \begin{bmatrix} 0 & -5 & 0 \\ 5 & 0 & 2 \\ 0 & -2 & 0 \end{bmatrix} = -\mathbf{B}$

(iii) If $\mathbf{C} = \begin{bmatrix} 0 & 1 & 7 & -4 \\ -1 & 0 & -4 & 3 \\ -7 & 4 & 0 & -2 \\ 4 & -3 & 2 & 0 \end{bmatrix}$, then $\mathbf{C}^{\mathrm{T}} = \begin{bmatrix} 0 & -1 & -7 & 4 \\ 1 & 0 & 4 & -3 \\ 7 & -4 & 0 & 2 \\ -4 & 3 & -2 & 0 \end{bmatrix} = -\mathbf{C}$.

Theorem 1.8 [Square Matrices]

Every square matrix $\mathbf{A}$ can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

Remarks: (i) For the square matrix $\mathbf{A}$, $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is a symmetric matrix and $\mathbf{A} - \mathbf{A}^{\mathrm{T}}$ is a skew-symmetric matrix.

(ii) Theorem 1.8 can be rewritten in the following form, that is, the square matrix $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \frac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} + \frac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2}.$$

EXAMPLE 1.11 Express the square matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$ as the sum of a symmetric matrix and a skew-symmetric matrix.

Solution: For the given matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$,

The transpose of **A** is $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}$

The symmetric matrix is $\dfrac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} = \begin{bmatrix} 1 & \frac{7}{2} \\ \frac{7}{2} & 4 \end{bmatrix}$.

The skew-symmetric matrix is $\dfrac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2} = \begin{bmatrix} 0 & \frac{-3}{2} \\ \frac{3}{2} & 0 \end{bmatrix}$.

The sum of the above symmetric matrix and the skew-symmetric matrix is

$$\frac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} + \frac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} = \mathbf{A}$$

Therefore, the given matrix **A** can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

EXAMPLE 1.12 Express the square matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \\ 6 & -8 & 2 \end{bmatrix}$ as the sum of a symmetric matrix and a skew-symmetric matrix.

Solution: For the given matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \\ 6 & -8 & 2 \end{bmatrix}$,

The transpose of **A** is $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 6 \\ -1 & 2 & -8 \\ 0 & 4 & 2 \end{bmatrix}$

The symmetric matrix is $\dfrac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & -2 \\ 3 & -2 & 2 \end{bmatrix}$

The skew-symmetric matrix is $\dfrac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2} = \begin{bmatrix} 0 & -1 & -3 \\ 1 & 0 & 6 \\ 3 & -6 & 0 \end{bmatrix}$

The sum of the above symmetric matrix and the skew-symmetric matrix is

$$\frac{\mathbf{A}+\mathbf{A}^{\mathrm{T}}}{2}+\frac{\mathbf{A}-\mathbf{A}^{\mathrm{T}}}{2}=\begin{bmatrix}1 & -1 & 0\\ 1 & 2 & 4\\ 6 & -8 & 2\end{bmatrix}=\mathbf{A}$$

Therefore, the given matrix **A** can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

Upper Triangular Matrix: A square matrix is called *upper triangular* if all the entries below the main diagonal are zero and those above it may or may not be zero. For example, the following matrices are upper triangular matrices. See the zeros enclosed in triangles.

$$\mathbf{A}=\begin{bmatrix}1 & 3\\ 0 & 2\end{bmatrix}; \qquad \mathbf{B}=\begin{bmatrix}3 & -1 & 5\\ 0 & 2 & 0\\ 0 & 0 & 4\end{bmatrix}$$

Lower Triangular Matrix: A square matrix is called *lower triangular* if all the entries above the main diagonal are zero and those below it may or may not be zero. For example, the following matrices are lower triangular matrices. See the zeros enclosed in triangles.

$$\mathbf{A}=\begin{bmatrix}7 & 0\\ 2 & 4\end{bmatrix}; \qquad \mathbf{B}=\begin{bmatrix}1 & 0 & 0\\ -6 & 2 & 0\\ 8 & 5 & 4\end{bmatrix}$$

Orthogonal Matrix: A square matrix **A** is called *orthogonal* if $\mathbf{A}\mathbf{A}^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}\mathbf{A}=\mathbf{I}$.

EXAMPLE 1.13 Show that $\mathbf{A}=\begin{bmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{bmatrix}$ is orthogonal.

Solution: First, we find the transpose of **A**,

$$\mathbf{A}^{\mathrm{T}}=\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\ 0 & 1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix}$$

$$\mathbf{A}\mathbf{A}^{\mathrm{T}}=\begin{bmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\ 0 & 1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix}$$

$$=\begin{bmatrix}\cos^2\theta+\sin^2\theta & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \sin^2\theta+\cos^2\theta\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$=\mathbf{I}$$

Therefore, the given matrix **A** is an orthogonal matrix.

EXAMPLE 1.14 If **A** and **B** are orthogonal matrices, then **AB** is also an orthogonal matrix.

Solution: Since **A** and **B** are orthogonal matrices,

$$\mathbf{AA}^{\mathrm{T}} = \mathbf{I} \quad \text{and} \quad \mathbf{BB}^{\mathrm{T}} = \mathbf{I}$$

Now,
$$\begin{aligned}(\mathbf{AB})\,(\mathbf{AB})^{\mathrm{T}} &= (\mathbf{AB})\,(\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}) = \mathbf{A}(\mathbf{BB}^{\mathrm{T}})\mathbf{A}^{\mathrm{T}}\\ &= \mathbf{AIA}^{\mathrm{T}} = (\mathbf{AI})\mathbf{A}^{\mathrm{T}}\\ &= \mathbf{AA}^{\mathrm{T}} = \mathbf{I}\end{aligned}$$

Therefore, **AB** is an orthogonal matrix.

Submatrix: A submatrix of **A** is any matrix obtained by deleting rows and columns of **A**.

Note that the matrix **A** itself is also considered to be a submatrix of **A**. For example, consider the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 6 & 1 & 9 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & -1 & -6 \\ 1 & 5 & 7 & 0 \end{bmatrix}$$

Note that the submatrix of **A** of order 4×4 is only one, that is, it is the **A** itself. Some other submatrices of **A** are given below:

$$\begin{bmatrix} 6 & 1 \\ 0 & -1 \end{bmatrix}; \quad \begin{bmatrix} 5 & 6 & 1 \\ 2 & 2 & 4 \\ 3 & 0 & -1 \end{bmatrix}; \quad \begin{bmatrix} 6 & 1 & 9 \\ 0 & -1 & -6 \end{bmatrix}; \quad \begin{bmatrix} 6 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} 5 & 7 \end{bmatrix}.$$

EXERCISE SET 2

1. If $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 2 \\ 4 & 5 & 0 \end{bmatrix}$, $\quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$, verify that $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}} \cdot \mathbf{A}^{\mathrm{T}}$.

2. If $\mathbf{A} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\quad \mathbf{B} = \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}$, $\quad \mathbf{C} = \begin{bmatrix} 3 & 0 \\ -2 & 1 \\ 0 & 1 \end{bmatrix}$, verify that $(\mathbf{ABC})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$.

3. Which of the following matrices are symmetric matrices:

   (i) $\begin{bmatrix} 1 & -2 & 3 \\ -2 & 7 & 0 \\ 3 & 0 & 6 \end{bmatrix}$ $\qquad$ (ii) $\begin{bmatrix} -1 & 6 \\ -6 & 2 \end{bmatrix}$ $\qquad$ (iii) $\begin{bmatrix} 0 & 5 & 6 & -9 \\ 5 & 7 & 2 & -3 \\ 6 & 2 & 0 & \frac{1}{2} \\ -9 & -3 & \frac{1}{2} & 6 \end{bmatrix}$.

4. Which of the following matrices are skew-symmetric matrices?

(i) $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 5 & 3 \\ -5 & 0 & -4 \\ -3 & 4 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 0 & 6 & -4 \\ -6 & 0 & 2 \\ 4 & -2 & 3 \end{bmatrix}$

5. Prove that if **A** and **B** are symmetric matrices, then (**A** + **B**) is also symmetric.

6. Express a square matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \\ 6 & -8 & 2 \end{bmatrix}$ as a sum of a symmetric matrix and a skew-symmetric matrix.

7. Show that $\mathbf{A} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$ is orthogonal.

1.3 SQUARE MATRICES

Definition: Inverse of Square Matrix

If **A** is a square matrix and if we can find another matrix of the same size, say **B**, such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}$$

then we call **A** *invertible* and we say that **B** is an *inverse* of the matrix **A**.

Note: The inverse of matrix **A** can be denoted by $\mathbf{A}^{-1}$, that is, $\mathbf{B} = \mathbf{A}^{-1}$.

Theorem 1.9 [Uniqueness of Inverse]

Suppose that matrix **A** is invertible and that both matrices **B** and **C** are inverses of **A**. Then **B** = **C**.

or

If **A** is invertible, then the inverse of **A** is unique.

EXAMPLE 1.15 Is $\mathbf{B} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ an inverse of $\mathbf{A} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$?

Solution:

$$\mathbf{AB} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}$$

$\therefore$ $\mathbf{AB} = \mathbf{I}$.

Thus, **B** is an inverse of **A**.

Theorem 1.10 [Properties of Inverse]

Suppose that **A** and **B** are invertible matrices of the same size. Then,

(i) **AB** is invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
(ii) $\mathbf{A}^{-1}$ is invertible and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

(iii) For $n = 0, 1, 2, \ldots, k$, $\mathbf{A}^n$ is invertible and $(\mathbf{A}^n)^{-1} = \mathbf{A}^{-n} = (\mathbf{A}^{-1})^n$.

(iv) If c is any nonzero scalar, then $c\mathbf{A}$ is invertible and $(c\mathbf{A})^{-1} = \dfrac{1}{c}\mathbf{A}^{-1}$.

(v) $\mathbf{A}^{\mathrm{T}}$ is invertible and $(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$.

Elementary Matrix

A square matrix is called an *elementary matrix* if it can be obtained by applying a single elementary row operation to the identity matrix of the same size.

EXAMPLE 1.16 Which of the following matrices are elementary matrices?

(i) $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ (iii) $\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$

(iv) $\mathbf{D} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (v) $\mathbf{E} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

Solution:

(i) $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R_2 \to (-4)R_2 \quad \sim \quad \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix} = \mathbf{A}.$

Here $\mathbf{A}$ is obtained by a single elementary row operation on $\mathbf{I}_2$. Hence $\mathbf{A}$ is an elementary matrix.

(ii) $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R_2 \to R_2 + 3R_1 \quad \sim \quad \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} = \mathbf{B}.$

Here $\mathbf{B}$ is obtained by a single elementary row operation on $\mathbf{I}_2$. Hence $\mathbf{B}$ is an elementary matrix.

(iii) $\mathbf{C}$ cannot be obtained by a single elementary row operation on $\mathbf{I}_2$. Hence $\mathbf{C}$ is not an elementary matrix.

(iv) $\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad R_1 \leftrightarrow R_3 \quad \sim \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \mathbf{D}$

Here $\mathbf{D}$ is obtained by a single elementary row operation on $\mathbf{I}_3$. Hence $\mathbf{D}$ is an elementary matrix.

(v) $\mathbf{E}$ cannot be obtained by a single elementary row operation on $\mathbf{I}_3$. Hence $\mathbf{E}$ is not an elementary matrix.

Row Operation by Matrix Multiplication

The following theorem suggests that the elementary row operation can be represented by the matrix multiplication.

Theorem 1.11 [Elementary Row Operation]

Suppose $\mathbf{E}$ is an elementary matrix that was found by applying an elementary row operation to $\mathbf{I}$. Then if $\mathbf{A}$ is an $n \times m$ matrix, $\mathbf{EA}$ is the matrix that will result by applying the same row operation to $\mathbf{A}$.

EXAMPLE 1.17 For the following matrix **A** perform the row operation $R_2 \to R_2 - 3R_1$ on it and then find the elementary matrix, **E**, for this operation and verify that **EA** will give the same result.

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 2 & 5 \\ 3 & 0 & 1 & -1 \\ 7 & 6 & 2 & 0 \end{bmatrix}$$

Solution: By applying the given row operation on **A**, we get

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 2 & 5 \\ 3 & 0 & 1 & -1 \\ 7 & 6 & 2 & 0 \end{bmatrix} \qquad R_2 \to R_2 - 3R_1 \qquad \sim \qquad \begin{bmatrix} 1 & 4 & 2 & 5 \\ 0 & -12 & -5 & -16 \\ 7 & 6 & 2 & 0 \end{bmatrix} = \mathbf{B}.$$

By applying the same row operation on $\mathbf{I}_3$, we get the elementary matrix **E**:

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_2 \to R_2 - 3R_1 \qquad \sim \qquad \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{E}.$$

We want to show that **EA** = **B**.

$$\mathbf{EA} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 2 & 5 \\ 3 & 0 & 1 & -1 \\ 7 & 6 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 2 & 5 \\ 0 & -12 & -5 & -16 \\ 7 & 6 & 2 & 0 \end{bmatrix} = \mathbf{B}.$$

EXAMPLE 1.18 Find the row operation that will take the elementary matrices **A**, **B**, **D** of Example 1.16 back to their respective original identity matrix.

Solution:

(i) $$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix} \qquad R_2 \to -\frac{1}{4} R_2 \qquad \sim \qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

(ii) $$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \qquad R_2 \to R_2 - 3R_1 \qquad \sim \qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

(iii) $$\mathbf{D} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad R_1 \leftrightarrow R_3 \qquad \sim \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}_3.$$

Note: These kinds of operations are called *inverse operations* and hence each row operation will have an inverse operation associated with it.

Theorem 1.12 [Inverse Operation]

Suppose that **E** is the elementary matrix associated with a particular row operation and that $\mathbf{E}_0$ is the elementary matrix associated with the inverse operation. Then **E** is invertible and $\mathbf{E}^{-1} = \mathbf{E}_0$.

Theorem 1.13 [Properties of Square Matrices]

If **A** is an $n \times n$ matrix, then the following statements are equivalent.

(i) **A** is invertible.

(ii) **A** is row equivalent to $\mathbf{I}_n$, that is, the reduced row-echelon form of **A** is $\mathbf{I}_n$.
(iii) **A** is expressible as a product of elementary matrices.

$$\mathbf{A} = \mathbf{E}_1^{-1}\,\mathbf{E}_2^{-1} \ldots \mathbf{E}_k^{-1}\mathbf{I}_n.$$

Method to Find the Inverse of a Square Matrix

From Theorem 1.11, Theorem 1.12 and the last result of Theorem 1.13, we can write

$$\mathbf{A}^{-1} = \mathbf{E}_1\,\mathbf{E}_2 \ldots \mathbf{E}_k\mathbf{I}_n.$$

If we consider each $\mathbf{E}_i$ as an elementary row operation, then the sequence of row operations that reduces **A** to $\mathbf{I}_n$, will reduce $\mathbf{I}_n$ to $\mathbf{A}^{-1}$:

$$\begin{aligned}(\mathbf{A} \vdots \mathbf{I}_n) &\to (\mathbf{E}_1\mathbf{A} \vdots \mathbf{E}_1\mathbf{I}_n)\\ &\to (\mathbf{E}_2\mathbf{E}_1\mathbf{A} \vdots \mathbf{E}_2\mathbf{E}_1\mathbf{I}_n)\\ &\to \ldots \quad \ldots \quad \ldots\\ &\to (\mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A} \vdots \mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{I}_n) = (\mathbf{I}_n \vdots \mathbf{A}^{-1})\end{aligned}$$

In other words, we consider an array with the matrix **A** on the left and the matrix $\mathbf{I}_n$ on the right. We now perform elementary row operations on the array and try to reduce the left-hand-half to the matrix $\mathbf{I}_n$. If we succeed in doing so, then the right-hand-half of the array gives the inverse $\mathbf{A}^{-1}$.

Thus, we have the following result:

> To find the inverse of an invertible matrix **A**, we must find a sequence of elementary row operations that reduces **A** to the identity matrix and then perform this same sequence of operation on $\mathbf{I}_n$ to obtain $\mathbf{A}^{-1}$.

EXAMPLE 1.19 Find the inverse of the following matrices:

(i) $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$ (iii) $\mathbf{C} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 4 & 5 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Solution:

(i) By writing the given matrix side by side with the unit matrix $\mathbf{I}_2$, we get

$$[\mathbf{A} \vdots \mathbf{I}_2] = \left[\begin{array}{cc:cc} 2 & 1 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$

$$R_2 \to R_2 - R_1 \quad \sim \left[\begin{array}{cc:cc} 2 & 1 & 1 & 0 \\ 1 & 3 & -1 & 1 \end{array}\right]$$

$$R_1 \leftrightarrow R_2 \quad \sim \left[\begin{array}{cc:cc} 1 & 3 & -1 & 1 \\ 2 & 1 & 1 & 0 \end{array}\right]$$

$$R_2 \to R_2 - 2R_1 \quad \sim \quad \begin{bmatrix} 1 & 3 & \vdots & -1 & 1 \\ 0 & -5 & \vdots & 3 & -2 \end{bmatrix}$$

$$R_2 \to -\frac{1}{5} R_2 \quad \sim \quad \begin{bmatrix} 1 & 3 & \vdots & -1 & 1 \\ 0 & 1 & \vdots & -\frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

$$R_1 \to R_1 - 3R_2 \quad \sim \quad \begin{bmatrix} 1 & 0 & \vdots & \frac{4}{5} & -\frac{1}{5} \\ 0 & 1 & \vdots & -\frac{3}{5} & \frac{2}{5} \end{bmatrix} = [\mathbf{I}_2 \vdots \mathbf{A}^{-1}].$$

Hence
$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix}.$$

(ii) By writing the given matrix side by side with the unit matrix $\mathbf{I}_3$, we get

$$[\mathbf{B} \vdots \mathbf{I}_3] = \begin{bmatrix} 0 & 1 & 2 & \vdots & 1 & 0 & 0 \\ 1 & 2 & 3 & \vdots & 0 & 1 & 0 \\ 3 & 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

$$R_1 \leftrightarrow R_2 \quad \sim \quad \begin{bmatrix} 1 & 2 & 3 & \vdots & 0 & 1 & 0 \\ 0 & 1 & 2 & \vdots & 1 & 0 & 0 \\ 3 & 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

$$R_3 \to R_3 - 3R_1 \quad \sim \quad \begin{bmatrix} 1 & 2 & 3 & \vdots & 0 & 1 & 0 \\ 0 & 1 & 2 & \vdots & 1 & 0 & 0 \\ 0 & -5 & -8 & \vdots & 0 & -3 & 1 \end{bmatrix}$$

$$R_1 \to R_1 - 2R_2,\ R_3 \to R_3 + 5R_2 \quad \sim \quad \begin{bmatrix} 1 & 0 & -1 & \vdots & -2 & 1 & 0 \\ 0 & 1 & 2 & \vdots & 1 & 0 & 0 \\ 0 & 0 & 2 & \vdots & 5 & -3 & 1 \end{bmatrix}$$

$$R_3 \to \frac{1}{2} R_3 \quad \sim \quad \begin{bmatrix} 1 & 0 & -1 & \vdots & -2 & 1 & 0 \\ 0 & 1 & 2 & \vdots & 1 & 0 & 0 \\ 0 & 0 & 1 & \vdots & \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

$$R_1 \to R_1 + R_3,\ R_2 \to R_2 - 2R_3 \quad \sim \quad \begin{bmatrix} 1 & 0 & 0 & \vdots & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \vdots & -4 & 3 & -1 \\ 0 & 0 & 1 & \vdots & \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix} = [\mathbf{I}_3 \vdots \mathbf{B}^{-1}]$$

Hence $$\mathbf{B}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}.$$

(iii) By writing the given matrix side by side with unit matrix $\mathbf{I}_4$, we get

$$[\mathbf{C} \vdots \mathbf{I}_4] = \left[\begin{array}{cccc:cccc} 1 & 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 2 & 2 & 4 & 5 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_2 \to R_2 - 2R_1 \quad \sim \left[\begin{array}{cccc:cccc} 1 & 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_4 \to R_4 + R_2 \quad \sim \left[\begin{array}{cccc:cccc} 1 & 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 & 0 & 1 \end{array}\right]$$

At this point, we observe that it is impossible to reduce the left-hand-half of the array to $\mathbf{I}_4$. Therefore, we cannot find the inverse using this method.

One question immediately arises in our mind. Does the inverse of a given matrix exist? What should be the criteria for invertibility? This will be discussed in the next section.

EXERCISE SET 3

1. Which of the following are elementary matrices?

(i) $\begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} -5 & 1 \\ 1 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (v) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

2. Find the inverse of the following matrices, if the same exists.

(i) $\begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix}$ (ii) $\begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}$ (iii) $\begin{bmatrix} 6 & -4 \\ -3 & 2 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$

3. If $\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$, prove that $\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}$.

4. If $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 9 & 3 \\ 1 & 4 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$, verify that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

5. If $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 5 & -2 \end{bmatrix}$, prove that $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

6. Show that $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & -\tan\frac{\theta}{2} \\ \tan\frac{\theta}{2} & 1 \end{bmatrix}\begin{bmatrix} 1 & \tan\frac{\theta}{2} \\ -\tan\frac{\theta}{2} & 1 \end{bmatrix}^{-1}$.

7. If $\mathbf{B} = \begin{bmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$, compute $\mathbf{B}^{-1}$ and prove that $\mathbf{B}^3 = \mathbf{B}^{-1}$.

1.4 DETERMINANTS OF $n \times n$ MATRICES

In Section 1.3, we discussed the question of the invertibility of a square matrix. In this section, we shall relate this question to the determinant of the matrix. As we shall see later, the task is reduced to checking whether this determinant is zero or nonzero. So, what is a determinant?

Let us start with a 1×1 matrix of the form

$$\mathbf{A} = [a].$$

Note that here $\mathbf{I} = [1]$. If $a \neq 0$, then clearly the matrix $\mathbf{A}$ is invertible, with inverse matrix

$$\mathbf{A}^{-1} = [a^{-1}].$$

On the other hand, if $a = 0$, then clearly no matrix, say $\mathbf{B}$, can satisfy $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, so the matrix $\mathbf{A}$ is not invertible. We, therefore, conclude that the value a is a good "determinant" to determine whether the 1×1 matrix $\mathbf{A}$ is invertible, since the matrix $\mathbf{A}$ is invertible, if and only if $a \neq 0$.

Let us then agree on the following definition.

Definition: Determinant of a* 1 × 1 *matrix

Suppose that $\mathbf{A} = [a]$ is a 1×1 matrix. We write

$$\det \mathbf{A} = a$$

and call this, the determinant of the 1×1 matrix $\mathbf{A}$.

Next, let us turn to 2×2 matrices of the form $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We shall use elementary row operations to find out whether this 2×2 matrix $\mathbf{A}$ is invertible. So, we consider the array

$$(\mathbf{A} : \mathbf{I}) = \begin{bmatrix} a & b & : & 1 & 0 \\ c & d & : & 0 & 1 \end{bmatrix} \tag{1.1}$$

and try to use elementary row operations to reduce the left-hand-half of the array (1.1) to a 2×2 identity matrix $\mathbf{I}$. Suppose first of all that $a = c = 0$. Then the array (1.1) becomes

$$\begin{bmatrix} 0 & b & : & 1 & 0 \\ 0 & d & : & 0 & 1 \end{bmatrix}$$

and hence it is impossible to reduce the left-hand-half of array (1.1) by elementary row operations to the matrix $\mathbf{I}$.

Consider next the case when $a \neq 0$. Multiplying row 2 of array (1.1) by a, we obtain

$$\begin{bmatrix} a & b & : & 1 & 0 \\ ac & ad & : & 0 & a \end{bmatrix}$$

Adding $(-c)$ times row 1 to row 2, we obtain

$$\begin{bmatrix} a & b & : & 1 & 0 \\ 0 & ad - bc & : & -c & a \end{bmatrix} \tag{1.2}$$

If $D = ad - bc = 0$, then the array (1.2) becomes

$$\begin{bmatrix} a & b & : & 1 & 0 \\ 0 & 0 & : & -c & a \end{bmatrix}$$

and hence it is impossible to reduce the left-hand-half of array (1.2) by elementary row operations to the 2×2 matrix $\mathbf{I}$. On the other hand, if $D = ad - bc \neq 0$, then the array (1.2) can be reduced by elementary row operations to

$$\begin{bmatrix} 1 & 0 & : & d/D & -b/D \\ 0 & 1 & : & -c/D & a/D \end{bmatrix}$$

so that

$$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ c & a \end{bmatrix}$$

Finally, the case when $c \neq 0$. Interchanging rows 1 and 2 of array (1.1), we obtain

$$\begin{bmatrix} c & d & : & 0 & 1 \\ a & b & : & 1 & 0 \end{bmatrix}$$

Multiplying row 2 of the array by c, we obtain

$$\begin{bmatrix} c & d & : & 0 & 1 \\ ac & bc & : & c & 0 \end{bmatrix}$$

Adding $-a$ times row 1 to row 2, we obtain

$$\begin{bmatrix} c & d & : & 0 & 1 \\ 0 & bc - ad & : & c & -a \end{bmatrix}$$

Multiplying row 2 by -1, we obtain

$$\begin{bmatrix} c & d & : & 0 & 1 \\ 0 & ad - bc & : & -c & a \end{bmatrix} \tag{1.3}$$

Again, if $D = ad - bc = 0$, then the array (1.3) becomes

$$\begin{bmatrix} c & d & : & 0 & 1 \\ 0 & 0 & : & -c & a \end{bmatrix}$$

and hence it is impossible to reduce the left-hand-half of array (1.3) by elementary row operations to the 2×2 matrix $\mathbf{I}$. On the other hand, if $D = ad - bc \neq 0$, then the array (1.3) can be reduced by elementary row operations to

$$\begin{bmatrix} 1 & 0 & : & d/D & -b/D \\ 0 & 1 & : & -c/D & a/D \end{bmatrix}$$

so that

$$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ c & a \end{bmatrix}$$

Finally, note that $a = c = 0$ is a special case of $ad - bc = 0$. We therefore conclude that the value $ad - bc$ is a good determinant to determine whether the 2×2 matrix $\mathbf{A}$ is invertible, since the matrix $\mathbf{A}$ is invertible, if and only if $ad - bc \neq 0$.

Let us then agree on the following definition.

Definition: Determinant of a 2 × 2 matrix

Suppose that $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a 2×2 matrix. We write

$$\det \mathbf{A} = ad - bc$$

and call this, the determinant of the matrix $\mathbf{A}$.

Determinants for Square Matrices of Higher Order

If we attempt to apply the argument for 2×2 matrices to 3×3 matrices or 4×4 matrices, it is very likely that we shall end up in a confusion with possibly no firm conclusion. Hence we will try a different approach.

Our approach is inductive in nature. In other words, we shall define the determinant of 2×2 matrices in terms of determinants of 1×1 matrices, define the determinant of 3×3 matrices in terms of determinants of 2×2 matrices, define the determinant of 4×4 matrices in terms of determinants of 3×3 matrices, and so on.

Suppose now that we have defined the determinant of $(n-1) \times (n-1)$ matrices. Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

be an $n \times n$ matrix. For every $i,\ j = 1, \ldots, n,$ let us delete row i and column j of the $n \times n$ matrix $\mathbf{A}$ to obtain the $(n-1) \times (n-1)$ matrix $\mathbf{B}$.

$$\mathbf{B} = \begin{bmatrix} a_{11} & \cdots & a_{1(j-1)} & \bullet & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & \bullet & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ \bullet & & \bullet & \bullet & \bullet & & \bullet \\ a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & \bullet & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & \cdots & a_{n(j-1)} & \bullet & a_{n(j+1)} & \cdots & a_{nn} \end{bmatrix}$$

Here $\bullet$ denotes that the entry has been deleted.

Definition: Cofactor

The number $C_{ij} = (-1)^{i+j} \det \mathbf{B}$ is called the *cofactor* of the entry a_{ij} of $\mathbf{A}$. In other words, the cofactor of the entry a_{ij} is obtained from $\mathbf{A}$ by first deleting the row and the column containing the entry a_{ij}, then calculating the determinant of the resulting $(n-1) \times (n-1)$ matrix, and finally, multiplying by a sign $(-1)^{i+j}$. Moreover, an $n \times n$ matrix $\mathbf{C} = [C_{ij}]$ is called the *cofactor matrix* of $\mathbf{A}$.

Note that the entries of $\mathbf{A}$ in row i are given by

$$(a_{i1}\ a_{i2}\ \ldots\ a_{in}).$$

Definition: Cofactor Expansion by Row

By the cofactor expansion of $\mathbf{A}$ by row i, we mean the expression

$$\sum_{j=1}^{n} a_{ij} C_{ij} = a_{i1} C_{i1} + \cdots + a_{in} C_{in} \tag{1.4}$$

Note that all the entries of $\mathbf{A}$ in column j are given by

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

Definition: Cofactor Expansion by Column

By the cofactor expansion of **A** by column j, we mean the expression

$$\sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + \cdots + a_{nj}C_{nj} \tag{1.5}$$

Theorem 1.14 [Cofactor Expansion]

Suppose that **A** is an $n \times n$ matrix. Then the expressions (1.4) and (1.5) are all equal and independent of the row or column chosen.

Definition: Determinant of Order n

Suppose that **A** is an $n \times n$ matrix. We call the common value in (1.4) and (1.5) the determinant of the matrix **A**, denoted by det **A**.

Let us check whether this agrees with our earlier definition of the determinant of a 2×2 matrix. Writing

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we have

$$C_{11} = a_{22} = d; \qquad C_{12} = -a_{21} = -c;$$
$$C_{21} = -a_{12} = -b; \qquad C_{22} = a_{11} = a$$

It follows that

by row 1: $a_{11}C_{11} + a_{12}C_{12} = ad - bc$

by row 2: $a_{21}C_{21} + a_{22}C_{22} = -bc + ad$

by column 1: $a_{11}C_{11} + a_{21}C_{21} = ad - bc$

by column 2: $a_{12}C_{12} + a_{22}C_{22} = -bc + ad$

The four values are clearly equal, and of the form $ad - bc$ as before.

EXAMPLE 1.20 Find the determinant of the following matrices:

(i) $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 2 & 3 & 0 & 5 \\ 1 & 4 & 0 & 2 \\ 5 & 4 & 8 & 5 \\ 2 & 1 & 0 & 5 \end{bmatrix}$ (iii) $\mathbf{D} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 4 & 5 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Solution:

(i) For the given matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$, cofactor expansion by row 1 is:

$$C_{11} = (-1)^{1+1} \det \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = (-1)^{1+1} (2 - 3) = -1$$

$$C_{12} = (-1)^{1+2} \det\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = (-1)^{1+2}\,(1-9) = 8$$

$$C_{13} = (-1)^{1+3} \det\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = (-1)^{1+3}\,(1-6) = -5$$

so that

$$\det \mathbf{A} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = 0(-1) + 1(8) + 2(-5) = -2$$

Alternatively, let us use the cofactor expansion by column 2. Then

$$C_{12} = (-1)^{1+2} \det\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = (-1)^{1+2}(1-9) = 8$$

$$C_{22} = (-1)^{2+2} \det\begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} = (-1)^{2+2}(0-6) = -6$$

$$C_{23} = (-1)^{2+3} \det\begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix} = (-1)^{2+3}(0-2) = 2$$

so that

$$\det \mathbf{A} = a_{12}C_{12} + a_{22}C_{22} + a_{23}C_{23} = 1(8) + 2(-6) + 1(2) = -2.$$

When using cofactor expansion, we should choose a row or column with as few nonzero entries as possible in order to minimize the calculations.

(ii) For the given matrix $\mathbf{B} = \begin{bmatrix} 2 & 3 & 0 & 5 \\ 1 & 4 & 0 & 2 \\ 5 & 4 & 8 & 5 \\ 2 & 1 & 0 & 5 \end{bmatrix}$, it is convenient to use cofactor expansion by column 3, and then

$$\begin{aligned} \det \mathbf{B} &= b_{13}C_{13} + b_{23}C_{23} + b_{33}C_{33} + b_{43}C_{43} \\ &= 8C_{33} \\ &= 8(-1)^{3+3} \det\begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 2 \\ 2 & 1 & 5 \end{bmatrix} = -16 \end{aligned}$$

(iii) For the given matrix $\mathbf{D} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 4 & 5 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, it is convenient to use cofactor expansion by row 4, and then

$$\begin{aligned} \det \mathbf{D} &= d_{41}C_{41} + d_{42}C_{42} + d_{43}C_{43} + d_{44}C_{44} \\ &= C_{44} \end{aligned}$$

$$= (-1)^{4+4} \det\begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 3 & 0 \end{bmatrix} = 1(0-12) - 1(0-0) + 2(6-0) = 0$$

Definition: Singular and Non-singular Matrices

A square matrix **A** is called a *singular matrix* if det **A** = 0. Otherwise **A** is called a *non-singular matrix*.

Some Important Results

(i) Suppose that a square matrix **A** has a zero row or has a zero column. Then det **A** = 0.

(ii) Suppose that **A** is an $n \times n$ triangular matrix. Then det **A** is the product of the diagonal entries, that is, det **A** = $a_{11}\, a_{22} \ldots a_{nn}$.

(iii) Suppose that **A** is an $n \times n$ matrix.

  (a) Suppose that the matrix **B** is obtained from the matrix **A** by interchanging two rows of **A**. Then det **B** = – det **A**.

  (b) Suppose that the matrix **B** is obtained from the matrix **A** by adding a multiple of one row of **A** to another row. Then det **B** = det **A**.

  (c) Suppose that the matrix **B** is obtained from the matrix **A** by multiplying one row of **A** by a nonzero constant c. Then det **B** = c det **A**.

(iv) For every $n \times n$ matrix **A**, we have det $\mathbf{A}^{\mathrm{T}}$ = det **A**.

(v) For every $n \times n$ matrices **A** and **B**, we have det **AB** = det **A** det **B**.

(vi) Suppose that the $n \times n$ matrix **A** is invertible. Then $\det \mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}}$.

Theorem 1.15 [Equivalent Statements]

If **A** is an $n \times n$ matrix, then the following statements are equivalent:

(i) The matrix **A** is invertible.

(ii) The matrices **A** and **I** are row equivalent.

(iii) The determinant det **A** $\neq 0$.

Definition: Adjoint of Matrix A

If **A** is an $n \times n$ matrix,

$$\operatorname{adj} \mathbf{A} = \mathbf{C}^{\mathrm{T}} = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$ where **C** is a cofactor matrix of **A**, is called the *adjoint* of the matrix **A**.

Theorem 1.16 [Invertible]

Suppose that an $n \times n$ matrix **A** is invertible. Then $\mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}$.

$$\text{adj } \mathbf{B} = \begin{bmatrix} C_{11} & C_{21} & C_{31} & C_{41} \\ C_{12} & C_{22} & C_{32} & C_{42} \\ C_{13} & C_{23} & C_{33} & C_{43} \\ C_{14} & C_{24} & C_{34} & C_{44} \end{bmatrix}$$

$$= \begin{bmatrix} \det\begin{bmatrix} 4 & 0 & 2 \\ 4 & 8 & 5 \\ 1 & 0 & 5 \end{bmatrix} & -\det\begin{bmatrix} 3 & 0 & 5 \\ 4 & 8 & 5 \\ 1 & 0 & 5 \end{bmatrix} & \det\begin{bmatrix} 3 & 0 & 5 \\ 4 & 0 & 2 \\ 1 & 0 & 5 \end{bmatrix} & -\det\begin{bmatrix} 3 & 0 & 5 \\ 4 & 0 & 2 \\ 4 & 8 & 5 \end{bmatrix} \\ -\det\begin{bmatrix} 1 & 0 & 2 \\ 5 & 8 & 5 \\ 2 & 0 & 5 \end{bmatrix} & \det\begin{bmatrix} 2 & 0 & 5 \\ 5 & 8 & 5 \\ 2 & 0 & 5 \end{bmatrix} & -\det\begin{bmatrix} 2 & 0 & 5 \\ 1 & 0 & 2 \\ 2 & 0 & 5 \end{bmatrix} & \det\begin{bmatrix} 2 & 0 & 5 \\ 1 & 0 & 2 \\ 5 & 8 & 5 \end{bmatrix} \\ \det\begin{bmatrix} 1 & 4 & 2 \\ 5 & 4 & 5 \\ 2 & 1 & 5 \end{bmatrix} & -\det\begin{bmatrix} 2 & 3 & 5 \\ 5 & 4 & 5 \\ 2 & 1 & 5 \end{bmatrix} & \det\begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 2 \\ 2 & 1 & 5 \end{bmatrix} & -\det\begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 2 \\ 5 & 4 & 5 \end{bmatrix} \\ -\det\begin{bmatrix} 1 & 4 & 0 \\ 5 & 4 & 8 \\ 2 & 1 & 0 \end{bmatrix} & \det\begin{bmatrix} 2 & 3 & 0 \\ 5 & 4 & 8 \\ 2 & 1 & 0 \end{bmatrix} & -\det\begin{bmatrix} 2 & 3 & 0 \\ 1 & 4 & 0 \\ 2 & 1 & 0 \end{bmatrix} & \det\begin{bmatrix} 2 & 3 & 0 \\ 1 & 4 & 0 \\ 5 & 4 & 8 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 144 & -80 & 0 & -112 \\ -8 & 0 & 0 & 8 \\ -51 & 30 & -2 & 41 \\ -56 & 32 & 0 & 40 \end{bmatrix}$$

By Theorem 1.16,

$$\mathbf{B}^{-1} = \frac{1}{\det \mathbf{B}} \text{adj } \mathbf{B}$$

$$= -\frac{1}{16}\begin{bmatrix} 144 & -80 & 0 & -112 \\ -8 & 0 & 0 & 8 \\ -51 & 30 & -2 & 41 \\ -56 & 32 & 0 & 40 \end{bmatrix}$$

$$= \begin{bmatrix} -9 & 5 & 0 & 7 \\ \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ \frac{51}{16} & -\frac{15}{8} & \frac{1}{8} & -\frac{41}{16} \\ \frac{7}{2} & -2 & 0 & -\frac{5}{2} \end{bmatrix}$$

$$\text{The third order minor } \det\left(\begin{bmatrix}1 & 4 & 5\\ 2 & 6 & 2\\ 3 & 7 & 1\end{bmatrix}\right) = \begin{vmatrix}1 & 4 & 5\\ 2 & 6 & 2\\ 3 & 7 & 1\end{vmatrix} = (-8+16-20) = -12.$$

$$\therefore \qquad \text{rank } \mathbf{D} = 3.$$

This method usually involves a lot of computational work to evaluate several determinants as we have seen in the above examples. This computational work can be reduced by using elementary row operations because the ranks of row equivalent matrices are equal. So we have the following method to determine the rank of a matrix.

Method 1.2: By the use of row-echelon form

Step 1: By the series of elementary (row or column or both) operations, convert an $m \times n$ matrix $\mathbf{A}$ into row echelon form (or upper triangular matrix).
Step 2: By the second definition of a rank, if the number of nonzero rows in the matrix obtained in Step 1 is r, then the rank $\mathbf{A} = r$.

EXAMPLE 1.25 Find the rank of the following matrices:

(i) $\mathbf{A} = \begin{bmatrix}1 & 2 & 1\\ -1 & 0 & 2\\ 2 & 3 & -3\end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix}1 & 2 & 3\\ 2 & 4 & 7\\ 3 & 6 & 10\end{bmatrix}$ (iii) $\mathbf{C} = \begin{bmatrix}1 & 2 & 3 & 0\\ 2 & 4 & 3 & 2\\ 3 & 2 & 1 & 3\\ 6 & 8 & 7 & 5\end{bmatrix}$

(iv) $\mathbf{D} = \begin{bmatrix}0 & 1 & 2 & -2\\ 4 & 0 & 2 & 6\\ 2 & 1 & 3 & 1\end{bmatrix}$

Solution:

(i) The given matrix $\mathbf{A} = \begin{bmatrix}1 & 2 & 1\\ -1 & 0 & 2\\ 2 & 3 & -3\end{bmatrix}$ is a 3×3 matrix. Therefore, rank $\mathbf{A} \le 3$.

Now we use the elementary row operation to find the row-echelon form of $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix}1 & 2 & 1\\ -1 & 0 & 2\\ 2 & 3 & -3\end{bmatrix}$$

$$R_2 \to R_2 + R_1,\ R_3 \to R_3 - 2R_1 \quad \sim \begin{bmatrix}1 & 2 & 1\\ 0 & 2 & 3\\ 0 & -1 & -5\end{bmatrix}$$

$$R_2 \to R_2 + 2R_3 \quad \sim \begin{bmatrix}1 & 2 & 1\\ 0 & 0 & -7\\ 0 & -1 & -5\end{bmatrix}$$

$$R_2 \leftrightarrow R_3 \quad \sim \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -5 \\ 0 & 0 & -7 \end{bmatrix}$$

$$R_2 \to (-1)\,R_2,\ R_3 \to -\frac{1}{7}\,R_3 \quad \sim \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$$

In the row-echelon form, **A** has 3 nonzero rows. Therefore, rank **A** = 3.

(ii) The given matrix $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 6 & 10 \end{bmatrix}$ is a 3×3 matrix. Therefore, rank $\mathbf{B} \leq 3$.

Now we use the elementary row operation to find the row-echelon form of **B**,

$$\mathbf{B} \quad = \quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 6 & 10 \end{bmatrix}$$

$$R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1 \quad \sim \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_3 \to R_3 - R_2 \quad \sim \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

In the row-echelon form, **B** has 2 nonzero rows. Therefore, rank **B** = 2.

(iii) The given matrix $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 6 & 8 & 7 & 5 \end{bmatrix}$ is a 4×4 matrix. Therefore, rank $\mathbf{C} \leq 4$.

Now we use the elementary row operation to find the row-echelon form of **C**.

$$\mathbf{C} \quad = \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 6 & 8 & 7 & 5 \end{bmatrix}$$

$$R_4 \to R_4 - (R_1 + R_2 + R_3) \quad \sim \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

For example, with x and y instead of x_1 and x_2, the linear equation $2x + 3y = 6$ describes the line passing through the points (3, 0) and (0, 2).

Similarly, with x, y and z instead of x_1, x_2 and x_3, the linear equation $2x + 3y + 4z = 12$ describes the plane passing through the points (6, 0, 0), (0, 4, 0), (0, 0, 3).

Definition: System of Linear Equations

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is a family of linear equations

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \tag{1.6}$$

We wish to determine if such a system has a solution, that is to find out if there exist numbers $x_1, x_2, \ldots, x_n$ which satisfy each of the equations simultaneously. We say that the system is *consistent* if it has a solution. Otherwise the system is called *inconsistent*.

Now we will see some examples of consistent and inconsistent system of two-dimensional space.

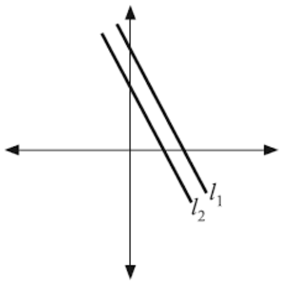


Figure 1.1 Parallel lines (no solution).

EXAMPLE 1.26 If you draw the two lines

$$2x + y = 4 \qquad \text{(line } l_1\text{)}$$

and

$$4x + 2y = 6 \qquad \text{(line } l_2\text{)}$$

it is easy to see that the two lines are parallel and hence do not intersect (Figure 1.1). Therefore, this system of two linear equations has no solution.

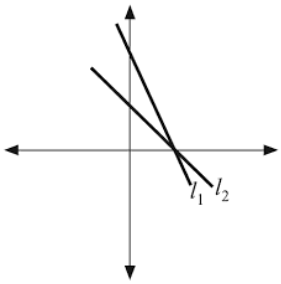


Figure 1.2 Intersecting lines (unique solution).

EXAMPLE 1.27 If you draw the two lines

$$2x + y = 4 \qquad \text{(line } l_1\text{)}$$

and

$$x + y = 2 \qquad \text{(line } l_2\text{)}$$

it is easy to see that the two lines are not parallel and intersect at the point (2, 0) (Figure 1.2). Hence this system of two linear equations has exactly one solution.

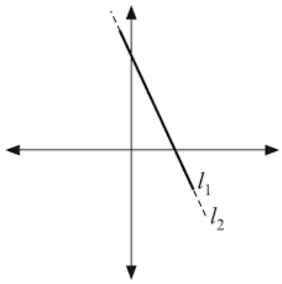


Figure 1.3 Overlapping lines (infinite solution).

EXAMPLE 1.28 If you draw the two lines

$$2x + y = 4 \qquad \text{(line } l_1\text{)}$$

and

$$4x + 2y = 8 \qquad \text{(line } l_2\text{)}$$

it is easy to see that the two lines overlap completely, hence this system of two linear equations has infinitely many solutions (Figure 1.3).

In the three examples above, we have shown that a system of two-linear equations on the plane R^2 may have no solution, one solution or infinitely many solutions. A natural question to ask is whether there can be any other conclusion. Well, we can see geometrically that two lines cannot intersect at more than one point without overlapping completely. Hence there can be no other conclusion.

Now if we consider the system of linear equations in n variables, then we may not be so lucky as to be able to see geometrically what is going on. We, therefore need to study the problem from a more algebraic viewpoint. So for the same purpose, first we start with the matrix form of the system of linear equations.

1.7 MATRIX FORM OF THE SYSTEM OF LINEAR EQUATIONS

The system of linear equations (1.6) can be written concisely as

$$\sum_{j=1}^{n} a_{ij}x_j = b_i; \qquad i = 1, 2, \ldots, m$$

If we consider the following matrices for the system (1.6),

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then the system (1.6) can be expressed as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

or

$$\mathbf{AX} = \mathbf{B}$$

which is called the matrix form of the system (1.6). Note that the matrix **A** is called *the coefficient matrix* for the system (1.6).

If we omit reference to the variables, then the system (1.6) can be represented by the array

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & : & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & : & b_2 \\ \vdots & \vdots & \cdots & \vdots & : & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & : & b_m \end{bmatrix}$$

of all the coefficients. This is known as the *augmented matrix* of the system. Here the first row of the array represents the first linear equation, and so on.

Let us go through some examples.

EXAMPLE 1.29 Find the matrix form of the following systems of linear equations.

(i)
$$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ x_1 + x_2 &= 4 \end{aligned}$$

(ii)
$$\begin{aligned} x_1 + 3x_2 - x_3 &= 10 \\ 4x_1 - x_2 + 7x_3 &= 15 \\ x_1 - 7x_3 &= 9 \end{aligned}$$

(iii)
$$\begin{aligned} x_1 + x_2 &= -5 \\ x_1 - x_2 - 6x_3 &= 6 \end{aligned}$$

Solution:

(i) For the given system,

$$\begin{aligned} 2x_1 - 3x_2 &= 5 \\ x_1 + x_2 &= 4 \end{aligned}$$

if we consider the following matrices

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

then the given system can be expressed in matrix form as

$$\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

or $$\mathbf{AX} = \mathbf{B}$$

(ii) For the given system,

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 10 \\ 4x_1 - x_2 + 7x_3 &= 15 \\ x_1 \qquad - 7x_3 &= 9 \end{aligned}$$

if we consider the following matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -1 \\ 4 & -1 & 7 \\ 1 & 0 & -7 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 10 \\ 15 \\ 9 \end{bmatrix}$$

then the given system can be expressed in matrix form as:

$$\begin{bmatrix} 1 & 3 & -1 \\ 4 & -1 & 7 \\ 1 & 0 & -7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 9 \end{bmatrix}$$

or $$\mathbf{AX} = \mathbf{B}$$

(iii) For the given system,

$$\begin{aligned} x_1 + x_2 \qquad &= -5 \\ x_1 - x_2 - 6x_3 &= 6 \end{aligned}$$

if we consider the following matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & -1 & -6 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} -5 \\ 6 \end{bmatrix}$$

then the given system can be expressed in matrix form as

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & -1 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ 6 \end{bmatrix}$$

or $$\mathbf{AX} = \mathbf{B}$$

The following theorem gives the condition for the consistency of the system (1.6).

Theorem 1.17 If $\mathbf{AX} = \mathbf{B}$ is the system of linear equations, then:

(i) It is *consistent* with *unique* solution if rank $\mathbf{A}$ = rank ($[\mathbf{A}:\mathbf{B}]$) = n, where n = number of unknowns

(ii) It is *consistent* with *infinite* solutions if rank $\mathbf{A}$ = rank ($[\mathbf{A}:\mathbf{B}]$) = r, $r < n$.

(iii) It is an *inconsistent* system (i.e. has *no solution*) if rank $\mathbf{A} \neq$ rank ($[\mathbf{A}:\mathbf{B}]$).

EXAMPLE 1.30 Determine which of the following systems are consistent?

(i) $$\begin{aligned} 2x_1 + 6x_2 &= -11 \\ 6x_1 + 20x_2 - 6x_3 &= -3 \\ 6x_2 - 18x_3 &= -1. \end{aligned}$$

(ii) $$\begin{aligned} x + y + z &= 3 \\ x + 2y + 3z &= 4 \\ x + 4y + 9z &= 6. \end{aligned}$$

(iii) $$\begin{aligned} x_1 + x_2 + x_3 - x_4 + x_5 &= 1 \\ 2x_1 - x_2 + 3x_3 + + 4x_5 &= 2 \\ 3x_1 - 2x_2 + 2x_3 + x_4 + x_5 &= 1 \\ x_1 + 0 + x_3 + 2x_4 + x_5 &= 0. \end{aligned}$$

Solution:

(i) The augmented matrix for the given system of linear equations is given by

$$[\mathbf{A}:\mathbf{B}] = \begin{bmatrix} 2 & 6 & 0 & : & -11 \\ 6 & 20 & -6 & : & -3 \\ 0 & 6 & -18 & : & -1 \end{bmatrix}$$

$$R_2 \to R_2 - 3R_1 \quad \sim \begin{bmatrix} 2 & 6 & 0 & : & -11 \\ 0 & 2 & -6 & : & 30 \\ 0 & 6 & -18 & : & -1 \end{bmatrix}$$

$$R_3 \to R_3 - 3R_2 \quad \sim \begin{bmatrix} 2 & 6 & 0 & : & -11 \\ 0 & 2 & -6 & : & 30 \\ 0 & 0 & 0 & : & -91 \end{bmatrix}$$

$$R_1 \to \frac{1}{2}R_1, R_2 \to \frac{1}{2}R_2 \quad \sim \begin{bmatrix} 1 & 3 & 0 & : & -\frac{11}{2} \\ 0 & 1 & -3 & : & 15 \\ 0 & 0 & 0 & : & -91 \end{bmatrix}$$

Therefore the rank $\mathbf{A}$ = 2 and rank ($[\mathbf{A}:\mathbf{B}]$) = 3, that is rank $\mathbf{A} \neq$ rank ($[\mathbf{A}:\mathbf{B}]$). So the given system of linear equation has no solution, that is, it is an inconsistent system.

(ii) The augmented matrix for the given system of linear equations is given by

$$[\mathbf{A}:\mathbf{B}] = \begin{bmatrix} 1 & 1 & 1 & : & 3 \\ 1 & 2 & 3 & : & 4 \\ 1 & 4 & 9 & : & 6 \end{bmatrix}$$

$$R_2 \to R_2 - R_1,\ R_3 \to R_3 - R_1 \quad \sim \begin{bmatrix} 1 & 1 & 1 & : & 3 \\ 0 & 1 & 2 & : & 1 \\ 0 & 3 & 8 & : & 3 \end{bmatrix}$$

$$R_3 \to R_3 - 3R_2 \quad \sim \begin{bmatrix} 1 & 1 & 1 & : & 3 \\ 0 & 1 & 2 & : & 1 \\ 0 & 0 & 2 & : & 0 \end{bmatrix}$$

$$R_3 \to \frac{1}{2} R_3 \quad \sim \begin{bmatrix} 1 & 1 & 1 & : & 3 \\ 0 & 1 & 2 & : & 1 \\ 0 & 0 & 1 & : & 0 \end{bmatrix}$$

Therefore the rank $\mathbf{A} = 3$ and rank $([\mathbf{A}:\mathbf{B}]) = 3$, that is

$$\text{rank } \mathbf{A} = \text{rank } ([\mathbf{A}:\mathbf{B}] = 3 = \text{number of unknown.}$$

So the given system of linear equation is a consistent system with unique solution.

(iii) The augmented matrix for the given system of linear equations is given by

$$[\mathbf{A}:\mathbf{B}] = \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 2 & -1 & 3 & 0 & 4 & : & 2 \\ 3 & -2 & 2 & 1 & 1 & : & 1 \\ 1 & 0 & 1 & 2 & 1 & : & 0 \end{bmatrix}$$

$$R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1,\ R_4 \to R_4 - R_1 \quad \sim \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 0 & -3 & 1 & 2 & 2 & : & 0 \\ 0 & -5 & -1 & 4 & -2 & : & -2 \\ 0 & -1 & 0 & 3 & 0 & : & -1 \end{bmatrix}$$

$$R_2 \leftrightarrow R_4 \quad \sim \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 0 & -1 & 0 & 3 & 0 & : & -1 \\ 0 & -5 & -1 & 4 & -2 & : & -2 \\ 0 & -3 & 1 & 2 & 2 & : & 0 \end{bmatrix}$$

$$R_3 \to R_3 - 5R_2,\ R_4 \to R_4 - 3R_2 \quad \sim \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 0 & -1 & 0 & 3 & 0 & : & -1 \\ 0 & 0 & -1 & -11 & -2 & : & 3 \\ 0 & 0 & 1 & -7 & 2 & : & 3 \end{bmatrix}$$

$$R_4 \to R_4 + R_3 \quad \sim \quad \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 0 & -1 & 0 & 3 & 0 & : & -1 \\ 0 & 0 & -1 & -11 & -2 & : & 3 \\ 0 & 0 & 0 & -18 & 0 & : & 6 \end{bmatrix}$$

$$R_2 \to (-1)\,R_2,\, R_3 \to (-1)\,R_3,\, R_4 \to \frac{1}{-18}R_4 \quad \sim \quad \begin{bmatrix} 1 & 1 & 1 & -1 & 1 & : & 1 \\ 0 & 1 & 0 & -3 & 0 & : & 1 \\ 0 & 0 & 1 & 11 & 2 & : & -3 \\ 0 & 0 & 0 & 1 & 0 & : & -\frac{1}{3} \end{bmatrix}$$

Therefore the rank $\mathbf{A} = 4$ and rank $([\mathbf{A:B}]) = 4$, that is

$$\text{rank } \mathbf{A} = \text{rank } ([\mathbf{A:B}] = 4 < 5 = \text{number of unknowns.}$$

So the given system of linear equations is a consistent system with infinitely many solutions.

EXAMPLE 1.31 Determine for what values of k the system of linear equations

$$\begin{aligned} 2x - 3y + 6z - 5t &= 3 \\ y - 4z + t &= 1 \\ 4x - 5y + 8z - 9t &= k \end{aligned}$$

has (i) no solution (ii) infinite solutions.

Solution: The augmented matrix for the given system of linear equations is given by

$$[\mathbf{A:B}] \quad = \quad \begin{bmatrix} 2 & -3 & 6 & -5 & : & 3 \\ 0 & 1 & -4 & 1 & : & 1 \\ 4 & -5 & 8 & -9 & : & k \end{bmatrix}$$

$$R_3 \to R_3 - 2R_1 \quad \sim \quad \begin{bmatrix} 2 & -3 & 6 & -5 & : & 3 \\ 0 & 1 & -4 & 1 & : & 1 \\ 0 & 1 & -4 & 1 & : & k-6 \end{bmatrix}$$

$$R_3 \to R_3 - R_2 \quad \sim \quad \begin{bmatrix} 2 & -3 & 6 & -5 & : & 3 \\ 0 & 1 & -4 & 1 & : & 1 \\ 0 & 0 & 0 & 0 & : & k-7 \end{bmatrix}$$

(i) If $k - 7 \neq 0$ or $k \neq 7$, then rank $\mathbf{A} = 2 \neq 3 =$ rank $([\mathbf{A:B}])$. So the system has no solution.

(ii) If $k - 7 = 0$ or $k = 7$, then

$$\text{rank } \mathbf{A} = \text{rank } ([\mathbf{A:B}] = 2 < 4 = \text{number of unknown}$$

So the system has infinite solutions.

EXAMPLE 1.32 Determine for what values of λ and μ the system of linear equations

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 10 \\ x + 2y + \lambda z &= \mu \end{aligned}$$

has (i) no solution (ii) a unique solution (iii) infinite solutions.

Solution: The augmented matrix for the given system of linear equations is given by

$$[\mathbf{A:B}] = \begin{bmatrix} 1 & 1 & 1 & : & 6 \\ 1 & 2 & 3 & : & 10 \\ 1 & 2 & \lambda & : & \mu \end{bmatrix}$$

$$R_3 \to R_3 - R_1 \quad \sim \begin{bmatrix} 1 & 1 & 1 & : & 6 \\ 0 & 1 & 2 & : & 4 \\ 0 & 1 & \lambda-1 & : & \mu-6 \end{bmatrix}$$

$$R_3 \to R_3 - R_2 \quad \sim \begin{bmatrix} 1 & 1 & 1 & : & 6 \\ 0 & 1 & 2 & : & 4 \\ 0 & 0 & \lambda-3 & : & \mu-10 \end{bmatrix}$$

(i) If $\lambda - 3 = 0$ or $\lambda = 3$ and $\mu - 10 \neq 0$ or $\mu \neq 10$, then rank $\mathbf{A} = 2 \neq 3 =$ rank $([\mathbf{A:B}])$. So the system has no solution.

(ii) If $\lambda - 3 \neq 0$ or $\lambda \neq 3$ and $\mu - 10 \neq 0$ or $\mu \neq 10$, then rank $\mathbf{A}$ = rank $([\mathbf{A:B}]) = 3$ = number of unknowns. So the system has a unique solution.

(iii) If $\lambda - 3 = 0$ or $\lambda = 3$ and $\mu - 10 = 0$ or $\mu = 10$, then

$$\text{rank } \mathbf{A} = \text{rank } ([\mathbf{A:B}]) = 2 < 3 = \text{number of unknowns.}$$

So the system has infinite solutions.

Definition: Homogeneous System

If $\mathbf{A}$ is an $m \times n$ matrix, then the system of linear equations of the form

$$\mathbf{AX} = \mathbf{0} \qquad (\text{i.e. } \mathbf{B} = \mathbf{0})$$

is called the *homogeneous system* of linear equations.

Note: The homogeneous system is a consistent system because it has always the trivial solution $\mathbf{X} = \mathbf{0}$, that is, $x_1 = x_2 = \cdots = x_n = 0$.

Theorem 1.18 [Homogeneous System]

If $\mathbf{A}$ is an $n \times n$ square matrix and $\mathbf{AX} = \mathbf{0}$ is the homogeneous system of linear equations, then

(i) It has a *unique trivial* solution $\mathbf{X} = \mathbf{0}$ if det $\mathbf{A} \neq 0$.

(ii) It has *non-trivial* solutions or *infinite* solutions if det $\mathbf{A} = 0$.

EXAMPLE 1.33 Find the values of k for which the following system of linear equations has non-trivial solutions:

$$3x + y - kz = 0$$
$$2x + 4y + kz = 0$$
$$8x - 4y - 6z = 0$$

Solution: The matrix form of the given system is

$$\mathbf{AX} = \mathbf{0}$$

where $\mathbf{A} = \begin{bmatrix} 3 & 1 & -k \\ 2 & 4 & k \\ 8 & -4 & -6 \end{bmatrix}$; $\quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$; $\quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Since **A** is a square matrix. So by Theorem 1.18, the system has non-trivial or infinite solutions if

$$\det \mathbf{A} = 0$$

$$\det\left(\begin{bmatrix} 3 & 1 & -k \\ 2 & 4 & k \\ 8 & -4 & -6 \end{bmatrix}\right) = 0$$

or $$3(-24 + 4k) - 1(-12 - 8k) - k(-8 - 32) = 0$$

or $$60k - 60 = 0$$

or $$k = 1$$

Therefore for $k = 1$, the system has non-trivial solutions or infinitely many solution.
From the above discussions, we have the following theorem.

Theorem 1.19 [Equivalent Statements]

If **A** is an $n \times n$ matrix, then the following statements are equivalent:

(i) $\det \mathbf{A} \neq 0$.
(ii) The reduced row-echelon form of **A** is **I**.
(iii) **A** is expressible as a product of elementary matrices.
(iv) **A** is invertible.
(v) $\mathbf{AX} = \mathbf{0}$ has only the trivial solution.
(vi) $\mathbf{AX} = \mathbf{B}$ is consistent for every $n \times 1$ matrix **B**.
(vii) $\mathbf{AX} = \mathbf{B}$ has exactly one solution for every $n \times 1$ matrix **B**.

In the next section, we will see some methods to determine the solutions of the system of linear equations.

EXERCISE SET 6

1. Find the matrix form of the following systems of linear equations.

(i) $$\begin{aligned} 2x + 3y &= 4 \\ 4x + 5y &= 6. \end{aligned}$$

(ii) $$\begin{aligned} x_1 &= 5 \\ x_2 &= 6. \end{aligned}$$

(iii) $$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ 2x_1 + x_2 \qquad &= 0 \\ 2x_2 + x_3 &= 1. \end{aligned}$$

(iv) $$\begin{aligned} x_1 + x_2 + \qquad x_4 &= 1 \\ x_1 + x_2 + x_3 \qquad &= 2 \\ x_2 + x_3 + x_4 &= 3 \\ x_1 \qquad\qquad + x_4 &= 0. \end{aligned}$$

2. Which of the following systems of linear equations are consistent?

(i) $$\begin{aligned} 2x + 3y - 4z &= 4 \\ x - y + 3z &= 4 \\ 3x + 2y - z &= -5. \end{aligned}$$

(ii) $$\begin{aligned} 3x + y + 2z &= 3 \\ 2x - 3y - z &= -3 \\ x + 2y + z &= 4. \end{aligned}$$

(iii) $$\begin{aligned} 5x + 3y + 7z &= 4 \\ 3x + 26y + 2z &= 9 \\ 7x + 2y + 10z &= 5. \end{aligned}$$

(iv) $$\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 6 \\ 6x_1 - 6x_2 + 6x_3 + 12x_4 &= 36 \\ 4x_1 + 3x_2 + 3x_3 - 3x_4 &= -1 \\ 2x_1 + 2x_2 - x_3 + x_4 &= 10. \end{aligned}$$

3. Find the values of λ for which the system

$$\begin{aligned} x + \ \ y + 4z &= 1 \\ x + 2y - 2z &= 1 \\ \lambda x + \ \ y + 2z &= 1 \end{aligned}$$

has (i) unique solution (ii) infinite solutions.

4. Determine for what values of a and b the system

$$\begin{aligned} x + \ \ y + 2z &= 6 \\ x + 2y + 4z &= 10 \\ x + \ \ y + az &= b \end{aligned}$$

has (i) no solution (ii) a unique solution.

5. Find the values of λ and μ for which the system

$$\begin{aligned} 2x + 3y + 5z &= 9 \\ 7x + 3y - 2z &= 8 \\ 2x + 3y + \lambda z &= \mu \end{aligned}$$

has (i) no solution (ii) a unique solution (iii) infinite solutions.

6. Find the values of a and b for which the system

$$\begin{aligned} x + 2y + 3z &= 6 \\ x + 3y + 5z &= 9 \\ 2x + 5y + az &= b \end{aligned}$$

has (i) no solution (ii) a unique solution (iii) infinite solutions.

7. Find the value of k such that the following systems of linear equations have non-trivial solutions.

(i) $$\begin{aligned} 4x_1 + 9x_2 + x_3 &= 0 \\ kx_1 + 3x_2 + kx_3 &= 0 \\ x_1 + 4x_2 + 2x_3 &= 0 \end{aligned}$$

(ii) $$\begin{aligned} 3x_1 + x_2 - kx_3 &= 0 \\ 2x_1 + 4x_2 + kx_3 &= 0 \\ 8x_1 - 4x_2 - 6x_3 &= 0 \end{aligned}$$

8. Show that the equations

$$\begin{aligned} 3x + 4y + 5z &= a \\ 4x + 5y + 6z &= b \\ 5x + 6y + 7z &= c \end{aligned}$$

do not have a solution unless $a + c = 2b$.

1.8 METHODS TO SOLVE THE SYSTEM OF LINEAR EQUATIONS

In the preceding sections, we saw that the system of linear equations has three possibilities for its solution: unique solution, infinite solutions or no solution. Now in this section, we will discuss the following four methods to determine the solutions of the system of the linear equations of the form **AX** = **B**:

(i) Inverse method (ii) Cramer's rule (iii) Gauss-elimination method and (iv) Gauss–Jordan method.

The first two methods (*Inverse method*, *Cramer's rule*) are applicable to only those systems whose coefficients matrix **A** is invertible and so it has a unique solution whereas the remaining two methods (*Gauss-elimination method*, *Gauss–Jordan method*) give the solutions of all the systems even if the coefficients matrix **A** is not invertible.

Inverse Method

Theorem 1.20 [System with Invertible Coefficients Matrix]
If **A** is an $n \times n$ invertible matrix, then the system of linear equation **AX** = **B** has exactly one solution, namely $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$.

Remark: This method is applicable only for that system whose coefficients matrix **A** is an invertible square matrix.

EXAMPLE 1.34 Solve the system of linear equations,

$$\begin{aligned} 2x_1 + x_2 &= 5 \\ 3x_1 + 4x_2 &= -10 \end{aligned}$$

with the help of matrix inverse.

Solution: The matrix form of the given system is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$; $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$

Since **A** is a square matrix with determinant det **A** = 5, so the inverse of matrix **A** exists and

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix}. \qquad \text{[see Example 1.19(i)]}$$

By Theorem 1.20, the solutions of the given system are given by the formula

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 5 \\ -10 \end{bmatrix}$$

namely,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$$

Therefore $x_1 = 6$ and $x_2 = -7$ is a solution of the given system.

EXAMPLE 1.35 Solve the system of linear equations,

$$\begin{aligned} y + 2z &= 2 \\ x + 2y + 3z &= 4 \\ 3x + y + z &= 6 \end{aligned}$$

with the help of matrix inverse.

Solution: The matrix form of the given system is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$; $\quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$; $\quad \mathbf{B} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$.

Since **A** is a square matrix with determinant det **A** = –2. So the inverse of matrix **A** exists and is given by

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$ [see Example 1.19(ii)]

By Theorem 1.20, the solutions of the given system are given by the formula

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 3 & -1 \\ \frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

i.e.
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

Therefore $x = 2$, $y = -2$ and $z = 2$ is a solution of the given system.

Remark: Consider a sequence of systems,

$$\mathbf{AX} = \mathbf{B}_1,\ \mathbf{AX} = \mathbf{B}_2,\ ...,\ \mathbf{AX} = \mathbf{B}_k,$$

each of which has the same square coefficient matrix **A**. If **A** is an invertible matrix, then by using the inverse method, the solutions can be obtained by the following formula:

$$\mathbf{X}_1 = \mathbf{A}^{-1}\mathbf{B}_1,\ \mathbf{X}_2 = \mathbf{A}^{-1}\mathbf{B}_2,\ ...,\ \mathbf{X}_k = \mathbf{A}^{-1}\mathbf{B}_k$$

EXAMPLE 1.36 Solve the sequence of systems

(a) $\begin{aligned} 3x_1 + x_2 &= 3 \\ -x_1 + 2x_2 + 2x_3 &= -3 \\ 5x_1 - x_3 &= 6 \cdot \end{aligned}$ $\qquad$ (b) $\begin{aligned} 3x_1 + x_2 &= -6 \\ -x_1 + 2x_2 + 2x_3 &= 12 \\ 5x_1 - x_3 &= 9 \cdot \end{aligned}$

Solution: The two systems have the same coefficient matrix $\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 2 & 2 \\ 5 & 0 & -1 \end{bmatrix}$ with different

$\mathbf{B}_1 = \begin{bmatrix} 3 \\ -3 \\ 6 \end{bmatrix}$ and $\mathbf{B}_2 = \begin{bmatrix} -6 \\ 12 \\ 9 \end{bmatrix}$.

Since the determinant det $\mathbf{A} = 3 \neq 0$, the inverse of matrix $\mathbf{A}$ exists and can be easily computed as

$$\mathbf{A}^{-1} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 3 & -1 & -2 \\ -\frac{10}{3} & \frac{5}{3} & \frac{7}{3} \end{bmatrix}$$

Now by applying the inverse method to both the systems, we get the solutions of both the systems as follows:

$$\mathbf{X}_1 = \mathbf{A}^{-1}\,\mathbf{B}_1 = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 3 & -1 & -2 \\ -\frac{10}{3} & \frac{5}{3} & \frac{7}{3} \end{bmatrix}\begin{bmatrix} 3 \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{X}_2 = \mathbf{A}^{-1}\,\mathbf{B}_2 = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 3 & -1 & -2 \\ -\frac{10}{3} & \frac{5}{3} & \frac{7}{3} \end{bmatrix}\begin{bmatrix} -6 \\ 12 \\ 9 \end{bmatrix} = \begin{bmatrix} 14 \\ -48 \\ 61 \end{bmatrix}$$

Therefore $x_1 = 1$, $x_2 = 0$ and $x_3 = -1$ is a solution of the system (a) and $x_1 = 14$, $x_2 = -48$ and $x_3 = 61$ is a solution of the system (b).

EXERCISE SET 7

Solve the following systems of linear equations with the help of the matrix inverse method.

1. $\begin{aligned} 2x - y + z &= 4 \\ x + y + z &= 1 \\ x - 3y - 2z &= 2. \end{aligned}$

2. $\begin{aligned} x + 2y + 3z &= 1 \\ 2x + 3y + 2z &= 2 \\ 3x + 3y + 4z &= 1. \end{aligned}$

3. $\begin{aligned} 2x - 2y + z &= 1 \\ x + 2y + 2z &= 2 \\ 2x + y - 2z &= 7. \end{aligned}$

4. $\begin{aligned} x + y + 4z &= 3 \\ x + 2y + 3z &= 4 \\ x + 4y + 9z &= 6. \end{aligned}$

5. $\begin{aligned} 5x + 3y + 3z &= 48 \\ 2x + 6y - 3z &= 18 \\ 8x - 3y + 2z &= 21. \end{aligned}$

6. $\begin{aligned} 4x + 4y + 3z &= -1 \\ 5x + y + 2z &= 1 \\ 7x + 3y + 4z &= 2. \end{aligned}$

7. $\begin{aligned} 7x_1 - 4x_2 \phantom{{}- 6x_3} &= 12 \\ -4x_1 + 12x_2 - 6x_3 &= 0 \\ -6x_2 + 14x_3 &= 0. \end{aligned}$

8. $\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 6 \\ 6x_1 - 6x_2 + 6x_3 + 12x_4 &= 36 \\ 4x_1 + 3x_2 + 3x_3 - 3x_4 &= -1 \\ 2x_1 + 2x_2 - x_3 + x_4 &= 10. \end{aligned}$

Cramer's Rule

Theorem 1.21 [Cramer's Rule]

Suppose that $\mathbf{A}$ is an $n \times n$ invertible matrix. Then the solution to the system $\mathrm{A}\mathbf{X} = \mathbf{B}$ is given by

$$x_1 = \frac{\det \mathbf{A}_1}{\det \mathbf{A}}, \quad x_2 = \frac{\det \mathbf{A}_2}{\det \mathbf{A}}, \ldots, x_n = \frac{\det \mathbf{A}_n}{\det \mathbf{A}}$$

where $\mathbf{A}_i$ is the matrix found by replacing the ith column of $\mathbf{A}$ with $\mathbf{B}$.

Remarks: For the following system of three unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

(i) If det $\mathbf{A} \neq 0$, then the system has a *unique* solution.

(ii) If det $\mathbf{A} = 0$ and at least one of three determinants det $\mathbf{A}_1$, det $\mathbf{A}_2$, det $\mathbf{A}_3$ is non-zero, then the system has no solution, that is, it is an *inconsistent* system.

(iii) If det $\mathbf{A} = 0$ and also det $\mathbf{A}_1 =$ det $\mathbf{A}_2 =$ det $\mathbf{A}_3 = 0$, then the system has an *infinite* number of solutions.

In general, for the system of n unknowns,

(iv) If det $\mathbf{A} \neq 0$, then the system has a *unique* solution.

(v) If det $\mathbf{A} = 0$ and at least one of det $\mathbf{A}_i$ $(1 \le i \le n)$ is nonzero, then the system has no solution, that is, it is an *inconsistent* system.

(vi) If det $\mathbf{A} = 0$ and also det $\mathbf{A}_1 =$ det $\mathbf{A}_2 = \cdots =$ det $\mathbf{A}_n = 0$, then the system has an *infinite* number of solutions.

EXAMPLE 1.37 Solve by Cramer's rule

$$\begin{aligned} x_1 - 3x_2 + x_3 &= 2 \\ 3x_1 + x_2 + x_3 &= 6 \\ 5x_1 + x_2 + 3x_3 &= 3\,. \end{aligned}$$

Solution: The matrix form of the given equation is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 1 & -3 & 1 \\ 3 & 1 & 1 \\ 5 & 1 & 3 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix}$; $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$$\det \mathbf{A} = \det\left(\begin{bmatrix} 1 & -3 & 1 \\ 3 & 1 & 1 \\ 5 & 1 & 3 \end{bmatrix}\right) = 1(3-1) + 3(9-5) + 1(3-5) = 12 \neq 0.$$

Since det $\mathbf{A} \neq 0$, the system has a *unique* solution. By Cramer's rule,

$$x_1 = \frac{\det \mathbf{A}_1}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 2 & -3 & 1 \\ 6 & 1 & 1 \\ 3 & 1 & 3 \end{bmatrix}\right)}{12} = \frac{52}{12} = \frac{13}{3}$$

$$x_2 = \frac{\det \mathbf{A}_2}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 1 \\ 5 & 3 & 3 \end{bmatrix}\right)}{12} = -\frac{14}{12} = -\frac{7}{6}$$

$$x_3 = \frac{\det \mathbf{A}_3}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 1 & -3 & 2 \\ 3 & 1 & 6 \\ 5 & 1 & 3 \end{bmatrix}\right)}{12} = -\frac{70}{12} = -\frac{35}{6}$$

EXAMPLE 1.38 Find the solution of the following system using Cramer's rule. Does the solution exist?

$$\begin{aligned} x_1 - x_2 \phantom{{}+2x_3} &= 1 \\ 2x_1 + x_2 + 2x_3 &= 2 \\ 3x_1 \phantom{{}+x_2} + 3x_3 &= 3 \end{aligned}$$

Solution: The matrix form of the given equation is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$; $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$$\det \mathbf{A} = \det\left(\begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{bmatrix}\right) = 1(3-0) + 1(6-6) + 0(0-3) = 3 \neq 0.$$

Since $\det \mathbf{A} \neq 0$, the system has a *unique* solution. By Cramer's rule,

$$x_1 = \frac{\det \mathbf{A}_1}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{bmatrix}\right)}{3} = \frac{3}{3} = 1$$

$$x_2 = \frac{\det \mathbf{A}_2}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}\right)}{3} = \frac{0}{3} = 0$$

$$x_3 = \frac{\det \mathbf{A}_3}{\det \mathbf{A}} = \frac{\det\left(\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{bmatrix}\right)}{3} = 0$$

EXAMPLE 1.39 Find the solution of the following system using Cramer's rule. Does the solution exist?

$$\begin{aligned} 2x - y - 6z &= 2 \\ -3x - 2y + z &= 0 \\ x + 3y + 5z &= 3 \end{aligned}$$

Solution: The matrix form of the given equation is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 2 & -1 & -6 \\ -3 & -2 & 1 \\ 1 & 3 & 5 \end{bmatrix}$; $\quad \mathbf{B} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$; $\quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$$\det \mathbf{A} = \det\left(\begin{bmatrix} 2 & -1 & -6 \\ -3 & -2 & 1 \\ 1 & 3 & 5 \end{bmatrix}\right) = 2(-10-3) + 1(-15-1) - 6(-9+2)$$

$$= -26 - 16 + 42 = 0.$$

$$\det \mathbf{A}_1 = \det\left(\begin{bmatrix} 2 & -1 & -6 \\ 0 & -2 & 1 \\ 3 & 3 & 5 \end{bmatrix}\right) = 2(-10-3) + 1\,(0-3) - 6(0+6) = -26 - 3 - 36$$

$$= -65 \neq 0$$

Since $\det \mathbf{A} = 0$ and $\det \mathbf{A}_1 \neq 0$, the system is inconsistent.

EXAMPLE 1.40 Find the solution of the following system using Cramer's rule. Does the solution exist?

$$\begin{aligned} 2x - y + 3z &= 4 \\ x + y - 3z &= -1 \\ 5x - y + 3z &= 7 \end{aligned}$$

Solution: The matrix form of the given equation is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & -3 \\ 5 & -1 & 3 \end{bmatrix}$; $\quad \mathbf{B} = \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix}$; $\quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

$$\det \mathbf{A} = \det\left(\begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & -3 \\ 5 & -1 & 3 \end{bmatrix}\right) = 0.$$

Similarly $\det \mathbf{A}_1 = \det \mathbf{A}_2 = \det \mathbf{A}_3 = 0$. Therefore, the system is consistent and it has infinite number of solutions.

Putting $z = k$ in the first two equations, we have

$$\begin{aligned} 2x - y &= 4 - 3k \\ x + y &= -1 + 3k \end{aligned}$$

By adding the two equations, we get

$$x = 1$$

Hence $x = 1$, $y = 3k - 2$ and $z = k$ is the solution for all values of k.

EXAMPLE 1.41 Does the homogeneous system

$$\begin{aligned} x + y - 3z &= 0 \\ 3x - y - z &= 0 \\ 2x + y - 4z &= 0 \end{aligned}$$

have non-trivial solutions? if so, find these solutions.

Solution: The matrix form of the given equation is

$$\mathbf{AX} = \mathbf{B}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$; $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

$$\det \mathbf{A} = \det\left(\begin{bmatrix} 1 & 1 & -3 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{bmatrix}\right) = 0.$$

Since det $\mathbf{A} = 0$, the system has non-trivial solutions.

To find these solutions, we solve any two of the given equations, say,

$$\begin{aligned} 3x - y - z &= 0 \\ 2x + y - 4z &= 0 \end{aligned}$$

We have
$$\frac{x}{\begin{vmatrix} -1 & -1 \\ 1 & -4 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} 3 & -1 \\ 2 & -4 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 3 & -1 \\ 2 & 1 \end{vmatrix}}$$

or
$$\frac{x}{5} = \frac{-y}{-10} = \frac{z}{5}$$

or
$$\frac{x}{1} = \frac{y}{2} = \frac{z}{1} = k$$

Hence $x = k$, $y = 2k$ and $z = k$ is the solution for all values of k.

EXAMPLE 1.42 Find the value of λ for which the equations

$$\begin{aligned} (2 - \lambda)x + 2y + 3 &= 0 \\ 2x + (4 - \lambda)y + 7 &= 0 \\ 2x + 5y + (6 - \lambda) &= 0 \end{aligned}$$

are consistent and find the values of x and y corresponding to each of these values of λ.

Solution: We are given a system of three equations in two unknowns which will be consistent, if

$$\det\left(\begin{bmatrix} 2-\lambda & 2 & 3 \\ 2 & 4-\lambda & 7 \\ 2 & 5 & 6-\lambda \end{bmatrix}\right) = 0$$

or
$$\det\left(\begin{bmatrix} 2-\lambda & 2 & 3 \\ 2 & 4-\lambda & 7 \\ 0 & 1+\lambda & -1-\lambda \end{bmatrix}\right) = 0$$

or
$$(\lambda + 1)(\lambda - 1)(\lambda - 12) = 0$$
$$\lambda = -1, 1, 12$$

***Case* (i)** when $\lambda = -1$, the first two equations become

$$3x + 2y + 3 = 0$$
$$2x + 5y + 7 = 0.$$

By using the elimination method, we get $x = -\dfrac{1}{11}, y = -\dfrac{15}{11}$.

***Case* (ii)** when $\lambda = 1$, the first two equations become

$$x + 2y + 3 = 0$$
$$2x + 3y + 7 = 0$$

By using the elimination method, we get $x = -5$, $y = 1$.

***Case* (iii)** when $\lambda = 12$, the first two equations become

$$-10x + 2y + 3 = 0$$
$$2x - 8y + 7 = 0.$$

By using the elimination method, we get $x = \dfrac{1}{2}, y = 1$.

EXERCISE SET 8

1. Find the solution of the following systems of linear equations using Cramer's rule. Does the solution exist in each case?

(i) $2x + y = 1$
$x - 2y = 8$

(ii) $x + y + z = 1$
$3x + 5y + 6z = 4$
$9x + 2y - 36z = 17.$

(iii) $2x - 3y + 7z = 5$
$x + y - 3z = 13$
$2x + 19y - 47z = 32.$

(iv) $2x - y + 3z = 4$
$x + y - 3z = -1$
$2x - y + 3z = 7.$

(v) $x + y + z = -1$
$x + 2y + 3z = -4$
$x + 3y + 4z = -6.$

(vi) $3x_1 + 2x_3 - x_4 = 60$
$2x_1 - x_2 + 4x_3 = 160$
$4x_2 + x_3 - 2x_4 = 20$
$5x_1 - x_2 - 2x_3 + x_4 = 10.$

2. Find the value of μ for which the following system of linear equation is consistent.

(i) $(\mu - 1)x_1 + (3\mu + 1)x_2 + 2\mu x_3 = 0$
$(\mu - 1)x_1 + (4\mu - 1)x_2 + (\mu + 3)x_3 = 0$
$2x_1 + (3\mu + 1)x_2 + 3(\mu - 1)x_3 = 0.$

(ii) $3x_1 - 2x_2 + 2x_3 = 3$
$x_1 + \mu x_2 - 3x_3 = 0$
$4x_1 + x_2 + 2x_3 = 7.$

3. Solve the following system of linear equation using Cramer's rule.

$$\begin{aligned} 2yz - 2x + xy &= 3xyz \\ 3yz - 2xz + 4xy &= 19xyz \\ 7zx - xy + 6yz &= 17xyz. \end{aligned}$$

4. Does the following system

$$\begin{aligned} 3x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 4x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \end{aligned}$$

have a non-trivial solution? If yes, find the solution.

5. Prove that the following system

$$\begin{aligned} x - ay - z &= 0 \\ y - z - ax &= 4 \\ z - x - y &= 0 \end{aligned}$$

is consistent when $a^3 + 1 = 0$.

6. If

$$\begin{aligned} bx &= ay - z \\ cy &= bz - x \\ az &= cx - y \end{aligned}$$

then prove that $a^2 + b^2 + c^2 = 0$.

Gauss-elimination Method

We now describe the Gauss elimination algorithm.

Step 1: Find the first nonzero column moving from left to right (column C_1) and select a nonzero entry from this column. By interchanging rows, if necessary, ensure that the first entry in this column is nonzero.

Step 2: Multiply row 1 by the multiplicative inverse $\dfrac{1}{a_{11}}$ of a_{11}, thereby converting a_{11} to 1.

Step 3: For each nonzero element a_{i1}, $i > 1$, (if any) in column C_1, add a_{i1} times row 1 to row i, thereby ensuring that all elements in column C_1, apart from the first, are zero.

Step 4: If the matrix obtained at Step 3 has its 2nd, …, mth rows all zero, the matrix is in row-echelon form. Otherwise suppose that the first column which has a nonzero element in the rows below the first row is column C_2. By interchanging rows below the first, if necessary, ensure that a_{22} is nonzero. Then convert a_{22} to 1 and by adding suitable multiples of row 2 to the rows below the row 2, where necessary, ensure that all elements below the row 2 in column C_2 are zero.

Continue in this way until the entire matrix is in *row-echelon form*.

Step 5: **(Back substitution)** To solve the corresponding system of equations, obtained from the row-echelon form, beginning with the bottom equation and working upwards, successively substitute each equation into all the equations above it.

Step 6: Assign arbitrary values to the free variables, if any.

EXAMPLE 1.43 Solve the following system of linear equations:

$$\begin{aligned} x + y + 2z &= 8 \\ -x - 2y + 3z &= 1 \\ 3x - 7y + 4z &= 10\cdot \end{aligned}$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ -1 & -2 & 3 \\ 3 & -7 & 4 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 8 \\ 1 \\ 10 \end{bmatrix}$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} 1 & 1 & 2 & : & 8 \\ -1 & -2 & 3 & : & 1 \\ 3 & -7 & 4 & : & 10 \end{bmatrix}$$

$$R_2 \to R_2 + R_1, R_3 \to R_3 - 3R_1 \quad \sim \begin{bmatrix} 1 & 1 & 2 & : & 8 \\ 0 & -1 & 5 & : & 9 \\ 0 & -10 & -2 & : & -14 \end{bmatrix}$$

$$R_3 \to R_3 - 10R_2, \quad \sim \begin{bmatrix} 1 & 1 & 2 & : & 8 \\ 0 & -1 & 5 & : & 9 \\ 0 & 0 & -52 & : & -104 \end{bmatrix}$$

$$R_3 \to -\frac{1}{52} R_3 \quad \sim \begin{bmatrix} 1 & 1 & 2 & : & 8 \\ 0 & -1 & 5 & : & 9 \\ 0 & 0 & 1 & : & 2 \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are:

$$\begin{aligned} x + y + 2z &= 8 \\ -y + 5z &= 9 \\ z &= 2. \end{aligned}$$

By substituting the value of z in the above two equations, we get

$$x = 3, y = 1, z = 2.$$

EXAMPLE 1.44 Solve the following system of linear equations:

$$\begin{aligned} 2x - 3y &= -2 \\ 2x + y &= 1 \\ 3x + 2y &= 1 \end{aligned}$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

The augmented matrix of the given system is

$$
\begin{aligned}
[\mathbf{A} : \mathbf{B}] &= \begin{bmatrix} 2 & -3 & : & -2 \\ 2 & 1 & : & 1 \\ 3 & 2 & : & 1 \end{bmatrix} \\
R_2 \to R_2 - R_1,\ R_3 \to R_3 - R_1 &\sim \begin{bmatrix} 2 & -3 & : & -2 \\ 0 & 4 & : & 3 \\ 1 & 5 & : & 3 \end{bmatrix} \\
R_1 \leftrightarrow R_3 &\sim \begin{bmatrix} 1 & 5 & : & 3 \\ 0 & 4 & : & 3 \\ 2 & -3 & : & -2 \end{bmatrix} \\
R_3 \to R_3 - 2R_1 &\sim \begin{bmatrix} 1 & 5 & : & 3 \\ 0 & 4 & : & 3 \\ 0 & -13 & : & -8 \end{bmatrix} \\
R_3 \to R_3 + 3R_2 &\sim \begin{bmatrix} 1 & 5 & : & 3 \\ 0 & 4 & : & 3 \\ 0 & -1 & : & 1 \end{bmatrix} \\
R_2 \to R_2 + 4R_3 &\sim \begin{bmatrix} 1 & 5 & : & 3 \\ 0 & 0 & : & 7 \\ 0 & -1 & : & 1 \end{bmatrix}
\end{aligned}
$$

The linear equations corresponding to the above row-echelon form are:

$$
\begin{aligned}
1x + 5y &= 3 \\
0x + 0y &= 7 \\
0x - \ \ y &= 1.
\end{aligned}
$$

Since the second equation cannot be satisfied, there is no solution to the system. Therefore the system is inconsistent.

EXAMPLE 1.45 Solve the following system of linear equations:

$$
\begin{aligned}
x - y + 2z - w &= -1 \\
2x - y + 2z - 2w &= -2 \\
-x + 2y - 4z + w &= 1 \\
3x \qquad\qquad - 3w &= -3.
\end{aligned}
$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$
\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 2 & -1 & 2 & -2 \\ -1 & 2 & -4 & 1 \\ 3 & 0 & 0 & -3 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} -1 \\ -2 \\ 1 \\ -3 \end{bmatrix}
$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} 1 & -1 & 2 & -1 & : & -1 \\ 2 & -1 & 2 & -2 & : & -2 \\ -1 & 2 & -4 & 1 & : & 1 \\ 3 & 0 & 0 & -3 & : & -3 \end{bmatrix}$$

$$R_2 \to R_2 - 2R_1,\ R_3 \to R_3 + R_1,\ R_4 \to R_4 - 3R_1 \quad \sim \begin{bmatrix} 1 & -1 & 2 & -1 & : & -1 \\ 0 & 1 & -2 & 0 & : & 0 \\ 0 & 1 & -2 & 0 & : & 0 \\ 0 & 3 & -6 & 0 & : & 0 \end{bmatrix}$$

$$R_3 \to R_3 - R_2,\ R_4 \to R_4 - 3R_2 \quad \sim \begin{bmatrix} 1 & -1 & 2 & -1 & : & -1 \\ 0 & 1 & -2 & 0 & : & 0 \\ 0 & 0 & 0 & 0 & : & 0 \\ 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are

$$x - y + 2z - w = -1$$
$$y - 2z = 0$$

Solving the equations for the leading variables,

$$x = y - 2z + w - 1$$
$$y = 2z$$

Substituting $y = 2z$ into the first equation, we get

$$x = w - 1$$
$$y = 2z$$

If we assign z and w the arbitrary values r and s, respectively, the general solution is given by the formulas

$$x = s - 1, \qquad y = 2r, \qquad z = r, \qquad w = s.$$

EXAMPLE 1.46 Solve the following system of linear equations:

$$\begin{aligned} 3x + 3y + 2z &= 1 \\ x + 2y &= 4 \\ 10y + 3z &= -2 \\ 2x - 3y - z &= 5. \end{aligned}$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 2 \\ 1 & 2 & 0 \\ 0 & 10 & 3 \\ 2 & -3 & -1 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 1 \\ 4 \\ -2 \\ 5 \end{bmatrix}.$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} 3 & 3 & 2 & : & 1 \\ 1 & 2 & 0 & : & 4 \\ 0 & 10 & 3 & : & -2 \\ 2 & -3 & -1 & : & 5 \end{bmatrix}$$

$$R_1 \leftrightarrow R_2 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 4 \\ 3 & 3 & 2 & : & 1 \\ 0 & 10 & 3 & : & -2 \\ 2 & -3 & -1 & : & 5 \end{bmatrix}$$

$$R_2 \to R_2 - 3R_1,\ R_4 \to R_4 - 2R_1 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 4 \\ 0 & -3 & 2 & : & -11 \\ 0 & 10 & 3 & : & -2 \\ 0 & -7 & -1 & : & -3 \end{bmatrix}$$

$$R_3 \to R_3 + 3R_2,\ R_4 \to R_4 - 2R_2 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 4 \\ 0 & -3 & 2 & : & -11 \\ 0 & 1 & 9 & : & -35 \\ 0 & -1 & -5 & : & 19 \end{bmatrix}$$

$$R_1 \to R_1 - 2R_3,\ R_2 \to R_2 + 3R_3,\ R_4 \to R_4 + R_3 \quad \sim \begin{bmatrix} 1 & 0 & -18 & : & 74 \\ 0 & 0 & 29 & : & -116 \\ 0 & 1 & 9 & : & -35 \\ 0 & 0 & 4 & : & -16 \end{bmatrix}$$

$$R_2 \leftrightarrow R_3,\ R_4 \to \frac{1}{4} R_4 \quad \sim \begin{bmatrix} 1 & 0 & -18 & : & 74 \\ 0 & 1 & 9 & : & -35 \\ 0 & 0 & 29 & : & -116 \\ 0 & 0 & 1 & : & -4 \end{bmatrix}$$

$$R_3 \to R_3 - 29R_4 \quad \sim \begin{bmatrix} 1 & 0 & -18 & : & 74 \\ 0 & 1 & 9 & : & -35 \\ 0 & 0 & 0 & : & 0 \\ 0 & 0 & 1 & : & -4 \end{bmatrix}$$

$$R_3 \leftrightarrow R_4 \quad \sim \begin{bmatrix} 1 & 0 & -18 & : & 74 \\ 0 & 1 & 9 & : & -35 \\ 0 & 0 & 1 & : & -4 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are:

$$\begin{aligned} x - 18z &= 74 \\ y + 9z &= -35 \\ z &= -4. \end{aligned}$$

By substituting the value of z in the above two equations, we get

$$x = 2,\ y = 1,\ z = -4$$

EXAMPLE 1.47 Solve the following homogeneous system of linear equations.

$$\begin{aligned} 2x + y + 3z &= 0 \\ x + 2y &= 0 \\ y + z &= 0. \end{aligned}$$

Solution: The matrix form of the given system of linear equations is:

$$\mathbf{AX} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} 2 & 1 & 3 & : & 0 \\ 1 & 2 & 0 & : & 0 \\ 0 & 1 & 1 & : & 0 \end{bmatrix}$$

$$R_1 \leftrightarrow R_2 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 0 \\ 2 & 1 & 3 & : & 0 \\ 0 & 1 & 1 & : & 0 \end{bmatrix}$$

$$R_2 \to R_2 - 2R_1 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 0 \\ 0 & -3 & 3 & : & 0 \\ 0 & 1 & 1 & : & 0 \end{bmatrix}$$

$$R_2 \to -\frac{1}{3} R_2 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 0 \\ 0 & 1 & -1 & : & 0 \\ 0 & 1 & 1 & : & 0 \end{bmatrix}$$

$$R_3 \to R_3 - R_2 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 0 \\ 0 & 1 & -1 & : & 0 \\ 0 & 0 & 2 & : & 0 \end{bmatrix}$$

$$R_3 \to \frac{1}{2} R_3 \quad \sim \begin{bmatrix} 1 & 2 & 0 & : & 0 \\ 0 & 1 & -1 & : & 0 \\ 0 & 0 & 1 & : & 0 \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are

$$\begin{aligned} x + 2y &= 0 \\ y - z &= 0 \\ z &= 0 \end{aligned}$$

By substituting the value of z in the above two equations, we get

$$x = 0, y = 0, z = 0.$$

Hence the given homogeneous system has only trivial solutions.

EXAMPLE 1.48 Solve the following system of linear equations

$$\begin{aligned} 2x + 2y + 4z \qquad &= 0 \\ - y - 3z + w &= 0 \\ 3x + y + z + 2w &= 0 \\ x + 3y - 2z - 2w &= 0. \end{aligned}$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 4 & 0 \\ 0 & -1 & -3 & 1 \\ 3 & 1 & 1 & 2 \\ 1 & 3 & -2 & -2 \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}; \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] \quad = \quad \begin{bmatrix} 2 & 2 & 4 & 0 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 3 & 1 & 1 & 2 & : & 0 \\ 1 & 3 & -2 & -2 & : & 0 \end{bmatrix}$$

$$R_1 \leftrightarrow R_4 \quad \sim \quad \begin{bmatrix} 1 & 3 & -2 & -2 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 3 & 1 & 1 & 2 & : & 0 \\ 2 & 2 & 4 & 0 & : & 0 \end{bmatrix}$$

$$R_3 \to R_3 - 3R_1, R_4 \to R_4 - 2R_1 \quad \sim \quad \begin{bmatrix} 1 & 3 & -2 & -2 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 0 & -8 & 7 & 8 & : & 0 \\ 0 & -4 & 8 & 4 & : & 0 \end{bmatrix}$$

$$R_3 \to R_3 - 8R_2, R_4 \to R_4 - 4R_2 \quad \sim \quad \begin{bmatrix} 1 & 3 & -2 & -2 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 0 & 0 & 31 & 0 & : & 0 \\ 0 & 0 & 20 & 0 & : & 0 \end{bmatrix}$$

$$R_3 \to \frac{1}{31}R_3,\ R_4 \to \frac{1}{20}R_4 \quad \sim \quad \begin{bmatrix} 1 & 3 & -2 & -2 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 0 & 0 & 1 & 0 & : & 0 \\ 0 & 0 & 1 & 0 & : & 0 \end{bmatrix}$$

$$R_4 \to R_4 - R_3 \quad \sim \quad \begin{bmatrix} 1 & 3 & -2 & -2 & : & 0 \\ 0 & -1 & -3 & 1 & : & 0 \\ 0 & 0 & 1 & 0 & : & 0 \\ 0 & 0 & 0 & 0 & : & 0 \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are

$$\begin{aligned} x + 3y - 2z - 2w &= 0 \\ -\ y - 3z + \ w &= 0 \\ z \qquad &= 0. \end{aligned}$$

Substituting $z = 0$ into the first two equations, we get

$$\begin{aligned} x + 3y - 2w &= 0 \\ -\ y + \ w &= 0 \\ z &= 0. \end{aligned}$$

Solving the equations for the leading variables,

$$\begin{aligned} x &= -3y + 2w \\ y &= w \\ z &= 0. \end{aligned}$$

Substituting $y = w$ into the first equation, we get

$$\begin{aligned} x &= -w \\ y &= w \\ z &= 0. \end{aligned}$$

If we assign w the arbitrary value t, the general solution is given by the formulas

$$x = -t, \qquad y = t, \qquad z = 0, \qquad w = t.$$

EXAMPLE 1.49 Solve the following system for x, y and z.

$$\begin{aligned} \frac{1}{x} + \frac{2}{y} - \frac{4}{z} &= 1 \\ \frac{2}{x} + \frac{3}{y} + \frac{8}{z} &= 0 \\ -\frac{1}{x} + \frac{9}{y} + \frac{10}{z} &= 5. \end{aligned}$$

Solution: If we consider $\frac{1}{x} = u, \frac{1}{y} = v$ and $\frac{1}{z} = w,$ the given system can be written as

$$\begin{aligned} u + 2v - 4w &= 1 \\ 2u + 3v + 8w &= 0 \\ -u + 9v + 10w &= 5. \end{aligned}$$

The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -4 \\ 2 & 3 & 8 \\ -1 & 9 & 10 \end{bmatrix}; \qquad \mathbf{X} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}.$$

The augmented matrix of the given system is

$$[\mathbf{A} : \mathbf{B}] = \begin{bmatrix} 1 & 2 & -4 & : & 1 \\ 2 & 3 & 8 & : & 0 \\ -1 & 9 & 10 & : & 5 \end{bmatrix}$$

$$R_1 \to R_2 - 2R_1,\ R_3 \to R_3 + R_1 \quad \sim \begin{bmatrix} 1 & 2 & -4 & : & 1 \\ 0 & -1 & 16 & : & -2 \\ 0 & 11 & 6 & : & 6 \end{bmatrix}$$

$$R_3 \to R_3 + 11R_2 \quad \sim \begin{bmatrix} 1 & 2 & -4 & : & 1 \\ 0 & -1 & 16 & : & -2 \\ 0 & 0 & 182 & : & -16 \end{bmatrix}$$

$$R_2 \to (-1)\,R_2,\ R_3 \to \frac{1}{182}\,R_3 \quad \sim \begin{bmatrix} 1 & 2 & -4 & : & 1 \\ 0 & 1 & -16 & : & 2 \\ 0 & 0 & 1 & : & -\frac{8}{91} \end{bmatrix}$$

The linear equations corresponding to the above row-echelon form are

$$u + 2v - 4w = 1$$
$$v - 16w = 2$$
$$w = -\frac{8}{91}.$$

By substituting the value of w in the above two equations, we get

$$u = -\frac{7}{13},\ v = \frac{54}{91},\ w = -\frac{8}{91}.$$

Therefore the solution in x, y and z variables is

$$x = \frac{1}{u} = -\frac{13}{7},\ y = \frac{1}{v} = \frac{91}{54},\ z = \frac{1}{w} = -\frac{91}{8}.$$

EXERCISE SET 9

1. Solve the following system of linear equations with the help of the matrix inverse method:

(i) $\begin{aligned} 2x + y &= 1 \\ x - 2y &= 8. \end{aligned}$

(ii) $\begin{aligned} 3x + y + 2z &= 3 \\ 2x - 3y - z &= -3 \\ x + 2y + z &= 4. \end{aligned}$

(iii) $\begin{aligned} 2x + y + 4z &= 2 \\ x + 3y - 2z &= 7 \\ 5x + 3y - 5z &= -8. \end{aligned}$

(iv) $\begin{aligned} 3x + y + z &= 8 \\ 2x - 3y - 2z &= -5 \\ 7x + 2y - 5z &= 0. \end{aligned}$

(v) $\begin{aligned} 3x + 3y + 2z &= 1 \\ x + 2y &= 4 \\ 10y + 3z &= 2. \end{aligned}$

(vi) $\begin{aligned} x + y - 3z &= 0 \\ 3x - y - z &= 0 \\ 2x + y - 4z &= 0. \end{aligned}$

(vii) $\begin{aligned} 3x - 2y + z + w &= 1 \\ x + y - 3z + 2w &= 2 \\ 6x + y + 4z + 3w &= 7. \end{aligned}$

(viii) $\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 6 \\ 4x_1 + 3x_2 + 3x_3 - 3x_4 &= -1 \\ 6x_1 - 6x_2 + 6x_3 + 12x_4 &= 36 \\ 2x_1 + 2x_2 - x_3 + x_4 &= 10. \end{aligned}$

Gauss–Jordan Elimination Method

Here we describe the Gauss–Jordan elimination algorithm.

Step 1: Find the first nonzero column moving from left to right, (column C_1) and select a nonzero entry from this column.

Step 2: By interchanging rows, if necessary; ensure that the first entry in this column is a nonzero. Multiply row 1 by the multiplicative inverse of a_{11}, thereby converting a_{11} to 1.

Step 3: For each nonzero element a_{i1}, $i > 1$, (if any) in column C_1, add a_{i1} times row 1 to row i, thereby ensuring that all elements in column C_1, apart from the first, are zero.

Step 4: If the matrix obtained at Step 3 has its 2nd, …, mth rows all zero, the matrix is in reduced row-echelon form. Otherwise suppose that the first column which has nonzero elements in the rows below the first is column C_2. By interchanging rows below the first, if necessary, ensure that a_{22} is nonzero. Then convert a_{22} to 1 and by adding suitable multiples of row 2 to the remaining rows, where necessary, ensure that all remaining elements in column C_2 are zero.

Continue in this way until the entire matrix is in *reduced row-echelon form*.

Step 5: To solve the corresponding system of equations, obtained from the reduced row-echelon form, first solve it for the leading variables and then assign arbitrary values to the free variables, if any.

EXAMPLE 1.50 Solve the following system of linear equations.

$$\begin{aligned} 3x + 2y - z &= -15 \\ 5x + 3y + 2z &= 0 \\ 3x + y + 3z &= 11 \\ -6x - 4y + 2z &= 30 \end{aligned}$$

Solution: The matrix form of the given system of linear equations is

$$\mathbf{AX} = \mathbf{B}$$

where

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & -1 \\ 5 & 3 & 2 \\ 3 & 1 & 3 \\ -6 & -4 & 2 \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} -15 \\ 0 \\ 11 \\ 30 \end{bmatrix}$$

The linear equations corresponding to the above reduced row-echelon form are

$$x - \frac{7}{2}z + \frac{5}{2}w = 0$$
$$y + 3z - 2w = 0.$$

Solving the equations for the leading variables,

$$x = \frac{7}{2}z - \frac{5}{2}w$$
$$y = -3z + 2w.$$

If we assign z and w the arbitrary values $2r$ and $2s$, respectively, the general solution is given by the formulas

$$x = 7r - 5s, \quad y = -6r + 4s, \quad z = 2r, \quad w = 2s.$$

EXERCISE SET 10

1. Solve the following system of linear equations with the help of the matrix inverse method.

(i) $\begin{aligned} 5x + 2y &= 3 \\ 3x + 2y &= 5. \end{aligned}$

(ii) $\begin{aligned} x + y + z &= 9 \\ 2x - 3y + 4z &= 13 \\ 3x + 4y + 5z &= 40. \end{aligned}$

(iii) $\begin{aligned} x + 3y + 6z &= 2 \\ 3x - y + 4z &= 9 \\ x - 4y + 2z &= 7. \end{aligned}$

(iv) $\begin{aligned} 3x + 3y + 2z &= 1 \\ x + 2y &= 4 \\ 10y + 3z &= -2 \\ 2x - 3y - z &= 5. \end{aligned}$

(v) $\begin{aligned} 7x - 2y &= -7 \\ 2x - y &= 1 \end{aligned}$

(vi) $\begin{aligned} 3x - 2y + z &= 1 \\ 5x + 3y + 3z &= 2 \\ 7x + 4y + 5z &= 3 \\ x + y - z &= 0 \end{aligned}$

(vii) $\begin{aligned} 3x + 2y + z &= 1 \\ 5x + 3y + 3z &= 2 \\ x + y - z &= 1. \end{aligned}$

(viii) $\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 6 \\ 4x_1 + 3x_2 + 3x_3 - 3x_4 &= -1 \\ 6x_1 - 6x_2 + 6x_3 + 12x_4 &= 36 \\ 2x_1 + 2x_2 - x_3 + x_4 &= 10. \end{aligned}$

(ix) $\begin{aligned} 2x - 3y &= -2 \\ 2x + y &= 1 \\ 3x + 2y &= 1. \end{aligned}$

(x) $\begin{aligned} 2x + y + 3z &= 0 \\ x + 2y &= 0 \\ y + z &= 0. \end{aligned}$

(xi) $\begin{aligned} x + 3y - 2z &= 5 \\ 2x + y + 4z &= 8 \\ 6x + y - 3z &= 5. \end{aligned}$

SUMMARY

Matrix A rectangular array of numbers of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is called an $m \times n$ matrix, with m rows and n columns. Here $m \times n$ is called the *order* of the matrix.

Theorem [Square Matrices] Every square matrix **A** can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

Remarks The above theorem can be rewritten in the following form, that is, the square matrix **A** can be expressed as: $\mathbf{A} = \dfrac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} + \dfrac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2}$, where $\mathbf{A} + \mathbf{A}^{\mathrm{T}}$ is the symmetric matrix and $\mathbf{A} - \mathbf{A}^{\mathrm{T}}$ is the skew-symmetric matrix.

Upper Triangular Matrix A square matrix is called *upper triangular* if all the entries below the main diagonal are zero and those above it may or may not be zero.

Lower Triangular Matrix A square matrix is called *lower triangular* if all the entries above the main diagonal are zero and those below it may or may not be zero.

Orthogonal Matrix A square matrix **A** is called *orthogonal* if $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$.

Inverse of a Square Matrix If **A** is a square matrix and if we can find another matrix of the same size, say **B**, such that

$$\mathbf{AB} = \mathbf{BA} = \mathrm{I}$$

then we call **A** *invertible* and we say that **B** is an *inverse* of the matrix **A**.

Note: The inverse of **A** can be denoted by $\mathbf{A}^{-1}$, i.e. $\mathbf{B} = \mathbf{A}^{-1}$.

Theorem [Uniqueness of Inverse] Suppose that **A** is invertible and that both **B** and **C** are inverses of **A**. Then **B** = **C**. If **A** is invertible, then the inverse of **A** is unique.

Theorem [Properties of Inverse] Suppose that **A** and **B** are invertible matrices of the same size. Then,

(i) **AB** is invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
(ii) $\mathbf{A}^{-1}$ is invertible and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
(iii) For $n = 0, 1, 2, \ldots, k$, $\mathbf{A}^n$ is invertible and $(\mathbf{A}^n)^{-1} = \mathbf{A}^{-n} = (\mathbf{A}^{-1})^n$.
(iv) If c is any nonzero scalar, then $c\mathbf{A}$ is invertible and $(c\mathbf{A})^{-1} = \dfrac{1}{c}\mathbf{A}^{-1}$.
(v) $\mathbf{A}^{\mathrm{T}}$ is invertible and $(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$.

Elementary Matrix A square matrix is called an *elementary matrix* if it can be obtained by applying a single elementary row operation to the identity matrix of the same size.

Theorem [Row Operation by Matrix Multiplications] Suppose **E** is an elementary matrix that was found by applying an elementary row operation to **I**. Then if **A** is an $n \times m$ matrix, **EA** is the matrix that will result by applying the same row operation to **A**.

Theorem [Inverse Operation] Suppose that **E** is the elementary matrix associated with a particular row operation and that $\mathbf{E}_0$ is the elementary matrix associated with the inverse operation. Then **E** is invertible and $\mathbf{E}^{-1} = \mathbf{E}_0$.

Theorem [Properties of Square Matrices] If **A** is an $n \times n$ matrix then the following statements are equivalent.

(i) **A** is invertible.
(ii) **A** is row equivalent to **I**, that is the reduced row-echelon form of **A** is **I**.
(iii) **A** is expressible as a product of elementary matrices.

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \ldots \mathbf{E}_k^{-1}\mathbf{I}.$$

(v) If det $\mathbf{A} = 0$ and at least one of det $\mathbf{A}_i$ ($1 \le i \le n$) is nonzero, then the system has no solution, that is, it is an *inconsistent* system.

(vi) If det $\mathbf{A} = 0$ and also det $\mathbf{A}_1 =$ det $\mathbf{A}_2 = \cdots =$ det $\mathbf{A}_n = 0$, then the system has an *infinite* number of solutions.

Algorithm for Gauss-elimination Method

Step 1 Find the first nonzero column moving from left to right, (column C_1) and select a nonzero entry from this column. By interchanging rows, if necessary, ensure that the first entry in this column is nonzero.

Step 2 Multiply row 1 by the multiplicative inverse $\dfrac{1}{a_{11}}$ of a_{11} thereby converting a_{11} to 1.

Step 3 For each nonzero element a_{i1}, $i > 1$, (if any) in column C_1, add a_{i1} times row 1 to row i, thereby ensuring that all elements in column C_1, apart from the first, are zero.

Step 4 If the matrix obtained at Step 3 has its 2nd, ... , mth rows all zero, the matrix is in row-echelon form. Otherwise suppose that the first column which has a nonzero element in the rows below the first row is column C_2. By interchanging rows below the first, if necessary, ensure that a_{22} is nonzero. Then convert a_{22} to 1 and by adding suitable multiples of row 2 to the rows below the row 2, where necessary, ensure that all elements below the row 2 in column C_2 are zero.

Continue in this way until the entire matrix is in *row-echelon form*.

Step 5 (Back substitution) To solve the corresponding system of equations, obtained from the row-echelon form, beginning with the bottom equation and working upwards, successively substitute each equation into all the equations above it.

Step 6 Assign arbitrary values to the free variables, if any.

Algorithm for Gauss–Jordan-elimination Method

Step 1 Find the first nonzero column moving from left to right, (column C_1) and select a nonzero entry from this column.

Step 2 By interchanging rows, if necessary; ensure that the first entry in this column is a nonzero. Multiply row 1 by the multiplicative inverse of a_{11}, thereby converting a_{11} to 1.

Step 3 For each nonzero element a_{i1}, $i > 1$, (if any) in column C_1, add a_{i1} times row 1 to row i, thereby ensuring that all elements in column C_1, apart from the first, are zero.

Step 4 If the matrix obtained at Step 3 has its 2nd, ... , mth rows all zero, the matrix is in reduced row-echelon form. Otherwise suppose that the first column which has a nonzero elements in the rows below the first is column C_2. By interchanging rows below the first, if necessary, ensure that a_{22} is nonzero. Then convert a_{22} to 1 and by adding suitable multiples of row 2 to the remaining rows, where necessary, ensure that all remaining elements in column C_2 are zero.

Continue in this way until the entire matrix is in *reduced row-echelon form*.

Step 5 To solve the corresponding system of equations, obtained from the reduced row-echelon form, first solve it for the leading variables and then assign arbitrary values to the free variables, if any.

2 Vector Spaces

2.1 VECTORS IN R^n

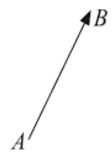


Figure 2.1 Geometric representation of a vector.

In this section, we will analyse the concept of a vector and hence the ultimate goal of this chapter will be to define something called Euclidean n-space. In the other sections, we will be looking at some very specific examples of vectors so that we can build up some of the ideas that surround them.

A vector is quantity that has magnitude and direction. A vector can be represented geometrically by a directed line segment that starts at point A, called the *initial point*, and ends at point B, called the *terminal point*. For instance, we could represent the vector in Figure 2.1 by $\mathbf{V} = \overrightarrow{AB}$.

Vectors in 2-Space and 3-Space

In 2-space, suppose that $\mathbf{v}$ is any vector whose initial point is at the origin of the rectangular co-ordinate system and its terminal point is at the co-ordinates (v_1, v_2) as shown in Figure 2.2. In this case, we call the co-ordinates of the terminal point, the components of the vector $\mathbf{v}$ and write,

$\mathbf{v} = [v_1, v_2]$ The square brackets will be used in the book to distinguish vector components from co-ordinates.

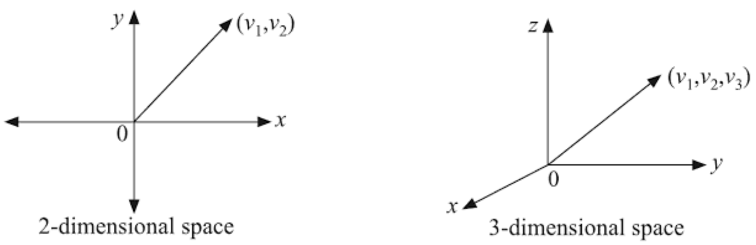


Figure 2.2 Vectors in R^2 and R^3.

We can do a similar thing for a vector **v** in 3-space; here v_1, v_2, v_3 are the components of the vector **v** and we write

$$\mathbf{v} = [v_1, v_2, v_3].$$

By generalizing, $[v_1, v_2, v_3, \ldots, v_n]$ are the components of vector **v** in R^n and we write

$$\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$$

Algebraic Operations on R^n

Now, we will define some algebraic operations like addition, scalar, multiplication, subtraction, etc. on the set of vectors.

Let us start with the definition of addition of two vectors.

Definition: Addition

If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ are two vectors in R^n then the addition of **u** and **v** is defined as

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots, u_n + v_n]$$

2-space

The addition of two vectors, $\mathbf{u} = [u_1, u_2]$ and $\mathbf{v} = [v_1, v_2]$ in R^2, (Figure 2.3) is defined as

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2]$$

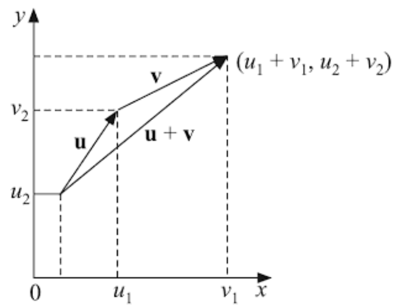


Figure 2.3 Vector addition in 2-dimensional space.

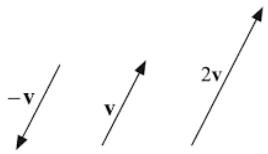
3-space

The addition of two vectors $\mathbf{u} = [u_1, u_2, u_3]$ and $\mathbf{v} = [v_1, v_2, v_3]$ in R^3, is defined as

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3]$$

Definition: Scalar Multiplication

If $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ is a vector and k is a nonzero scalar (i.e. k is any number), then the scalar multiple, $k\mathbf{v}$, is the vector whose length is k times the length of **v** and is in the direction of **v** if k is positive, and in opposite direction of **v** if k is negative. In other words,



$$k\mathbf{v} = [kv_1, kv_2, kv_3, \ldots, kv_n]$$

Subtraction

If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$, then the difference of **u** from **v**, denoted by $\mathbf{v} - \mathbf{u}$ (Figure 2.4) is defined to be

$$\begin{aligned}\mathbf{v} - \mathbf{u} &= \mathbf{v} + (-\mathbf{u})\\ &= [u_1 - v_1, u_2 - v_2, u_3 - v_3, \ldots, u_n - v_n]\end{aligned}$$

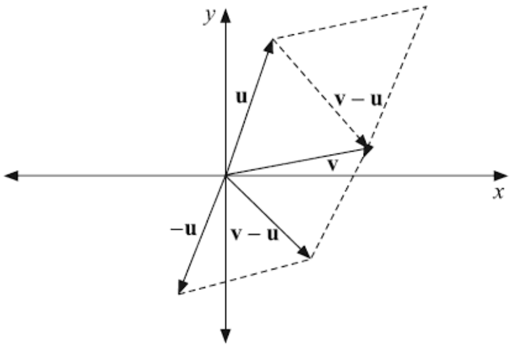


Figure 2.4 Vector subtraction in 2-dimensional space.

Remark: The zero vector in R^n is denoted by $\mathbf{0}$ and is defined to be the vector $\mathbf{0} = [0, 0, 0, \ldots, 0]$.

Theorem 2.1 [Properties of Vector in R^n]

If $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors in R^n and k and m are scalars, then

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
(ii) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
(iv) $\mathbf{u} - \mathbf{u} = \mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
(v) $1\mathbf{u} = \mathbf{u}$
(vi) $(km)\mathbf{u} = k(m\mathbf{u}) = m(k\mathbf{u})$
(vii) $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$
(viii) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

Norm and Distance

Definition: Norm of a Vector

If $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ is a vector in R^n, then the magnitude (Figure 2.5) of the vector is called the norm of the vector $\mathbf{v}$ and it is denoted by $\|\mathbf{v}\|$ and defined by the formula, $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

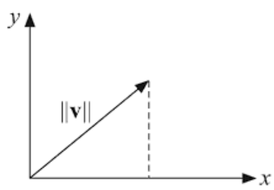


Figure 2.5

2-Space

In R^n, if $\mathbf{v} = [v_1, v_2]$ then $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$

Remark: We call $\mathbf{v}$ a unit vector if $\|\mathbf{v}\| = 1$. Given a nonzero vector $\mathbf{v}$ in R^n, defines a new vector $\mathbf{n} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$, then $\mathbf{n}$ is a unit vector.

Theorem 2.2 [Properties of Norm in R^n]

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in R^n and k be any scalar, then

(i) $\|\mathbf{u}\| \geq 0$, and $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
(ii) $\|k\mathbf{u}\| = |k|\,\|\mathbf{u}\|$
(iii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Definition: Distance between Two Vectors

The distance between two vectors, $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ in R^n is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

Theorem 2.3 [Properties of Distance in R^n]

Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be vectors in R^n and k be any scalar, then

(i) $d(\mathbf{u}, \mathbf{v}) \geq 0$, and $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$.
(ii) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
(iii) $d(\mathbf{u} + \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

EXAMPLE 2.1 Let $\mathbf{u} = [4, 1, 2, 3]$, $\mathbf{v} = [0, 3, 8, -2]$ and $\mathbf{w} = [3, 1, 2, 2]$. Evaluate each of the following expressions.

(i) $\|\mathbf{u} + \mathbf{v}\|$ (ii) $\|\mathbf{u}\| + \|\mathbf{v}\|$ (iii) $\dfrac{\mathbf{w}}{\|\mathbf{w}\|}$

(iv) $\left\|\dfrac{\mathbf{w}}{\|\mathbf{w}\|}\right\|$ (v) $\|4\mathbf{u}\| + 4\,\|\mathbf{u}\|$ (vi) $d(\mathbf{v}, \mathbf{u})$

Solution: Here $\mathbf{u} = [4, 1, 2, 3]$, $\mathbf{v} = [0, 3, 8, -2]$, $\mathbf{w} = [3, 1, 2, 2]$

(i) $\mathbf{u} + \mathbf{v} = [4, 1, 2, 3] + [0, 3, 8, -2] = [4 + 0, 1 + 3, 2 + 8, 3 - 2]$

$$= [4, 4, 10, 1]$$

$$\|\mathbf{u} + \mathbf{v}\| = \sqrt{4^2 + 4^2 + 10^2 + 1^2} = \sqrt{133}$$

(ii) $\|\mathbf{u}\| = \sqrt{4^2 + 1^2 + 2^2 + 3^2} = \sqrt{30}$

$$\|\mathbf{v}\| = \sqrt{0^2 + 3^2 + 8^2 + (-2)^2} = \sqrt{77}$$

$$\|\mathbf{u}\| + \|\mathbf{v}\| = \sqrt{30} + \sqrt{77}$$

(iii) $\|\mathbf{w}\| = \sqrt{3^2 + 1^2 + 2^2 + 2^2} = \sqrt{18} = 3\sqrt{2}$

$$\therefore \quad \frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{1}{3\sqrt{2}}[3, 1, 2, 2] = \left(\frac{1}{\sqrt{2}}, \frac{1}{3\sqrt{2}}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}\right)$$

(iv)
$$\left\|\frac{\mathbf{w}}{\|\mathbf{w}\|}\right\| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{3\sqrt{2}}\right)^2 + \left(\frac{\sqrt{2}}{3}\right)^2 + \left(\frac{\sqrt{2}}{3}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{18} + \frac{2}{9} + \frac{2}{9}} = 1.$$

(v) $\|4\mathbf{u}\| + 4\,\|\mathbf{u}\| = \|[16, 4, 8, 12]\| + 4\sqrt{30} = 4\sqrt{30} + 4\sqrt{30} = 8\sqrt{30}.$

(vi) $d(\mathbf{v}, \mathbf{u}) = \|\mathbf{v} - \mathbf{u}\|$

$$= \|[-4, 2, 6, -5]\|$$

$$= \sqrt{(-4)^2 + 2^2 + 6^2 + (-5)^2}$$

$$= \sqrt{81} = 9.$$

Definition: Dot Product in R^n

If **u** and **v** are two nonzero vectors in R^n and θ is the angle between them, then the dot product or the inner product $\mathbf{u} \cdot \mathbf{v}$ in R^n is the real number defined by

$$\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\ ||\mathbf{v}|| \cos\theta$$

Another definition of dot product in R^n in terms of components is as follows:

If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, , \ldots, v_n]$ are any vectors in R^n, then the dot-product or inner product $\mathbf{u} \cdot \mathbf{v}$ is defined by the formula,

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Remark: Another notation for the inner product is written as $\langle \mathbf{u}, \mathbf{v} \rangle$ which shall also be used in this book.

EXAMPLE 2.2 Find the values of the inner product $\mathbf{u} \cdot \mathbf{v}$, if:

(i) $\mathbf{u} = [2, 5]$, $\mathbf{v} = [-4, 3]$

(ii) $\mathbf{u} = [3, 1, 4, -5]$, $\mathbf{v} = [2, 2, -4, -3]$

Solution:

(i) The inner product in R^2 is given by the formula

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2$$

For the given vectors $\mathbf{u} = [2, 5]$, $\mathbf{v} = [-4, 3]$,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (2)(-4) + (5)(3) \\ &= -8 + 15 \\ &= 7 \end{aligned}$$

(ii) The inner product in R^4 is given by the formula,

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$$

For the given vectors $\mathbf{u} = [3, 1, 4, -5]$, $\mathbf{v} = [2, 2, -4, -3]$

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (3)(2) + (1)(2) + (4)(-4) + (-5)(-3) \\ &= 6 + 2 - 16 + 15 = 7 \end{aligned}$$

EXAMPLE 2.3 Find the vector in R^2 with norm 1 whose inner product with $(3, -1)$ is zero.

Solution: Let $\mathbf{u} = [u_1, u_2]$ and $\mathbf{v} = [3, -1]$

Suppose $||\mathbf{u}|| = 1$ and its inner product with **v** is zero, that is

$$||\mathbf{u}|| = 1 \quad \Rightarrow \quad u_1^2 + u_2^2 = 1 \qquad \text{(i)}$$

$$\mathbf{u} \cdot \mathbf{v} = 0 \quad \Rightarrow \quad 3u_1 - u_2 = 0 \qquad \text{(ii)}$$

Solving Eqs. (i) and (ii), we get

$$u_1 = \frac{1}{\sqrt{10}}, u_2 = \frac{3}{\sqrt{10}} \quad \text{or} \quad u_1 = \frac{-1}{\sqrt{10}}, u_2 = \frac{-3}{\sqrt{10}}$$

Therefore $\left[\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right]$ and $\left[-\frac{1}{\sqrt{10}}, -\frac{3}{\sqrt{10}}\right]$ are two vectors in R^2 with norm 1 whose inner products with $\mathbf{v} = [3, -1]$ are zero.

EXAMPLE 2.4 Find the angle between the vectors **u** and **v** shown below:

(i) $\mathbf{u} = [1, -1]$, $\mathbf{v} = [-1, 3]$ (ii) $\mathbf{u} = [2, -3, 1]$, $\mathbf{v} = [1, 1, 1]$

(iii) $\mathbf{u} = [1, 3, 1]$, $\mathbf{v} = [-3, 1, 1]$

Solution:

(i) Let θ be an angle between **u** and **v**, then $\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$

For the given vectors $\mathbf{u} = [1, -1]$, $\mathbf{v} = [-1, 3]$,

$$\mathbf{u}\cdot\mathbf{v} = [1, -1]\cdot[-1, 3] = -1 - 3 = -4$$

$$\|\mathbf{u}\| = \sqrt{1^2 + (-1)^2} = \sqrt{2}, \qquad \|\mathbf{v}\| = \sqrt{(-1)^2 + 3^2} = \sqrt{10}$$

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = -\frac{4}{\sqrt{2}\sqrt{10}} = -\frac{2}{\sqrt{5}}$$

$$\therefore \qquad \theta = \cos^{-1}\left(-\frac{2}{\sqrt{5}}\right)$$

(ii) For the given vectors $\mathbf{u} = [2, -3, 1]$, $\mathbf{v} = [1, 1, 1]$

$$\mathbf{u}\cdot\mathbf{v} = [2, -3, 1]\cdot[1, 1, 1] = 2 - 3 + 1 = 0$$

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = 0$$

$$\therefore \qquad \theta = \frac{\pi}{2}$$

(iii) For the given vectors $\mathbf{u} = [1, 3, 1]$, $\mathbf{v} = [-3, 1, 1]$

$$\mathbf{u}\cdot\mathbf{v} = [1, 3, 1]\cdot[-3, 1, 1] = -3 + 3 + 1 = 1$$

$$\|\mathbf{u}\| = \sqrt{1^2 + 3^2 + 1^2} = \sqrt{11}, \qquad \|\mathbf{v}\| = \sqrt{(-3)^2 + 1^2 + 1^2} = \sqrt{11}$$

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{1}{\sqrt{11}\sqrt{11}} = \frac{1}{11}$$

$$\therefore \qquad \theta = \cos^{-1}\left(\frac{1}{11}\right)$$

The following theorem lists the most important properties of the dot product.

Theorem 2.4 [Properties of the Dot Product]

Let **u**, **v**, **w** be the vectors in R^3 (or in R^2) and α, β be the real numbers. Then

(i) $\mathbf{u}\cdot\mathbf{u} = \|\mathbf{u}\|^2$. Hence $\mathbf{u}\cdot\mathbf{u} \geq 0$

(ii) $\mathbf{u}\cdot\mathbf{u} = 0$ if and if $\mathbf{u} = \mathbf{0}$.

(iii) $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$

(iv) $\mathbf{u}\cdot(\mathbf{v} + \mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$

(v) $(\alpha\mathbf{u})\cdot\mathbf{v} = \alpha(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(\alpha\mathbf{v})$

(vi) If θ is an angle between **u** and **v**, then

$$\theta \text{ is acute if and only if } \mathbf{u} \cdot \mathbf{v} > 0$$
$$\theta \text{ is obtuse if and only if } \mathbf{u} \cdot \mathbf{v} < 0$$
$$\theta = \frac{\pi}{2} \text{ if and only if } \mathbf{u} \cdot \mathbf{v} = 0.$$

The property (vi) in the preceding theorem gives us the following definition.

Definition: Orthogonal Vectors

Two vectors **u** and **v** in R^n are called orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

EXAMPLE 2.5 Determine whether **u** and **v** make an acute angle, an obtuse angle or are orthogonal

(i) $\mathbf{u} = [2, -2, 2]$, $\mathbf{v} = [0, 4, -2]$
(ii) $\mathbf{u} = [4, 8, 2]$, $\mathbf{v} = [0, 1, 3]$
(iii) $\mathbf{u} = [1, -2]$, $\mathbf{v} = [2, 1]$

Solution:

(i) Here $\mathbf{u} = [2, -2, 2]$, $\mathbf{v} = [0, 4, -2]$

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= [2, -2, 2] \cdot [0, 4, -2] \\ &= 0 - 8 - 4 = -10\end{aligned}$$

Since $\mathbf{u} \cdot \mathbf{v} < 0$, **u** and **v** make an obtuse angle (see Theorem 2.4).

(ii) For the given vectors $\mathbf{u} = [4, 8, 2]$, $\mathbf{v} = [0, 1, 3]$

$$\mathbf{u} \cdot \mathbf{v} = [4, 8, 2] \cdot [0, 1, 3] = 0 + 8 + 6 = 14$$
$$\mathbf{u} \cdot \mathbf{v} > 0$$

Since $\mathbf{u} \cdot \mathbf{v} > 0$, **u** and **v** make an acute angle (see Theorem 2.4).

(iii) For $\mathbf{u} = [1, -2]$, $\mathbf{v} = [2, 1]$

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= [1, -2] \cdot [2, 1] \\ &= 2 - 2 = 0\end{aligned}$$

Since $\mathbf{u} \cdot \mathbf{v} = 0$, **u** and **v** are orthogonal (see Theorem 2.4).

EXAMPLE 2.6 For which values of k are $\mathbf{u} = [k, k, 1]$ and $\mathbf{v} = [k, 5, 6]$ orthogonal?

Solution: By the definition, two vectors $\mathbf{u} = [k, k, 1]$ and $\mathbf{v} = [k, 5, 6]$ are orthogonal if

$$\mathbf{u} \cdot \mathbf{v} = 0$$

i.e. $$[k, k, 1] \cdot [k, 5, 6] = 0$$

or $$k^2 + 5k + 6 = 0$$

or $$(k + 2)(k + 3) = 0$$

$\therefore$ $$k = -2 \quad \text{or} \quad k = -3$$

EXAMPLE 2.7 Find a unit vector that is orthogonal to both $\mathbf{u} = [1, 0, 1]$ and $\mathbf{v} = [0, 1, 1]$.

Solution: Suppose $\mathbf{w} = [w_1, w_2, w_3]$ is a unit vector that is orthogonal to both $\mathbf{u} = [1, 0, 1]$ and $\mathbf{v} = [0, 1, 1]$.

That is
$$||\mathbf{w}|| = 1 \quad \Rightarrow \quad w_1^2 + w_2^2 + w_3^2 = 1$$
$$\mathbf{w} \cdot \mathbf{u} = 0 \quad \Rightarrow \quad w_1 + w_3 = 0$$
$$\mathbf{w} \cdot \mathbf{u} = 0 \quad \Rightarrow \quad w_2 + w_3 = 0$$

Solving the above equations, we get
$$w_1 = \frac{1}{\sqrt{3}},\ w_2 = \frac{1}{\sqrt{3}},\ w_3 = -\frac{1}{\sqrt{3}}$$

Hence $\mathbf{w} = \left[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right]$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

EXAMPLE 2.8 Find the angle between a diagonal of a cube and one of its edges.

Solution: We can see that, $\mathbf{v}_1 = [a, 0, 0]$ is one of its edges and $\mathbf{d} = [a, a, a]$ is a diagonal vector. Let θ be an angle between $\mathbf{v}_1$ and $\mathbf{d}$, then
$$\cos\theta = \frac{\mathbf{v}_1 \cdot \mathbf{d}}{\|\mathbf{v}_1\|\|\mathbf{d}\|} = \frac{[a, 0, 0] \cdot [a, a, a]}{(a)(\sqrt{3}a)}$$
$$= \frac{a^2}{\sqrt{3}a^2} = \frac{1}{\sqrt{3}}$$
$$\therefore \qquad \theta = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right)$$

One of the most important inequalities in linear algebra is the Cauchy–Schwarz inequality which has been described in the following theorem.

Theorem 2.5 [Cauchy–Schwarz Inequality in R^n]

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in R^n, then $|\mathbf{u} \cdot \mathbf{v}| \le ||\mathbf{u}||\ ||\mathbf{v}||$.

EXAMPLE 2.9 Use the Cauchy–Schwarz inequality to prove that for all real values of a, b, and θ
$$(a\cos\theta + b\sin\theta)^2 \le a^2 + b^2$$

Solution: Let $\mathbf{u} = [a, b]$ and $\mathbf{v} = [\cos\theta, \sin\theta]$
$$\mathbf{u} \cdot \mathbf{v} = [a, b] \cdot [\cos\theta, \sin\theta]$$
$$= a\cos\theta + b\sin\theta$$
$$||\mathbf{u}||^2 = a^2 + b^2; \quad ||\mathbf{v}||^2 = \cos^2\theta + \sin^2\theta$$

By the Cauchy–Schwarz inequality,
$$|\mathbf{u} \cdot \mathbf{v}| \le ||\mathbf{u}||\ ||\mathbf{v}||$$
$$[\mathbf{u} \cdot \mathbf{v}]^2 \le ||\mathbf{u}||^2\ ||\mathbf{v}||^2 \qquad \text{(By squaring both sides)}$$
$$\therefore \qquad (a\cos\theta + b\sin\theta)^2 \le a^2 + b^2$$

EXAMPLE 2.10 Verify Cauchy–Schwarz inequality for the vectors $\mathbf{u} = [0, -2, 2, 1]$, $\mathbf{v} = [-1, -1, 1, 1]$.

Solution: For $\mathbf{u} = [0, -2, 2, 1]$, $v = [-1, -1, 1, 1]$

$$\mathbf{u} \cdot \mathbf{v} = [0, -2, 2, 1] \cdot [-1, -1, 1, 1]$$

$$= 0 + 2 + 2 + 1 = 5$$

$$|\mathbf{u} \cdot \mathbf{v}| = 5$$

$$||\mathbf{u}|| = \sqrt{0^2 + (-2)^2 + 2^2 + 1^2} = 3$$

$$||\mathbf{v}|| = \sqrt{(-1)^2 + (-1)^2 + 1^2 + 1^2} = 2$$

$$||\mathbf{u}||\,||\mathbf{v}|| = 6$$

$$\therefore \qquad |\mathbf{u} \cdot \mathbf{v}| \le ||\mathbf{u}||\,||\mathbf{v}|| \qquad (\text{since } 5 < 6)$$

The following theorem shows an important property of the orthogonal vectors.

Theorem 2.6 [Pythagoras Theorem]
If **u** and **v** are orthogonal vectors in R^n, then $||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$

EXAMPLE 2.11 Verify Pythagoras Theorem for the vectors $\mathbf{u} = [1, -2, 1]$ and $\mathbf{v} = [2, 1, 0]$.

Solution: For $\mathbf{u} = [1, -2, 1]$, $\mathbf{v} = [2, 1, 0]$

$$\mathbf{u} \cdot \mathbf{v} = [1, -2, 1] \cdot [2, 1, 0]$$

$$= 2 - 2 = 0$$

Therefore **u** and **v** are orthogonal vectors.

$$||\mathbf{u}|| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}$$

$$||\mathbf{v}|| = \sqrt{2^2 + 1^2 + 0^2} = \sqrt{5}$$

$$\mathbf{u} + \mathbf{v} = [3, -1, 1] \quad \Rightarrow \quad ||u + v|| = \sqrt{3^2 + (-1)^2 + 1^2} = \sqrt{11}$$

$$||\mathbf{u}||^2 + ||\mathbf{v}||^2 = 6 + 5 = 11$$

Hence $$||\mathbf{u}||^2 + ||\mathbf{v}||^2 = ||\mathbf{u} + \mathbf{v}||^2$$

The Dot Product in the Form of Matrix Multiplication

Sometimes a vector $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ in R^n can be considered as a row matrix or a column matrix, that is,

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 & u_3 & \ldots & u_n \end{bmatrix} \qquad \text{or} \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

In this book, this vector is considered as the column matrix. One can easily check that the operations like addition, scalar multiplication, etc. on vectors are same as those on matrices.

$$\mathbf{u}+\mathbf{v} = \alpha\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix}$$

$$\alpha\mathbf{u} = \alpha\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{bmatrix}$$

Now we will try to find the matrix form of the dot product operation.

Let $\mathbf{u} = [u_1\ u_2\ u_3\ \ldots,\ u_n]$ and $\mathbf{v} = [v_1\ v_2\ v_3\ \ldots,\ v_n]$ be two vectors in R^n, then the co-ordinates form of the dot product is

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

If we consider $\mathbf{u}$ and $\mathbf{v}$ as column vectors, $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, then

$$\begin{aligned} \mathbf{v}^{\mathrm{T}}\mathbf{u} &= [v_1\ v_2\ v_3\ \ldots\ v_n]\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \\ &= [u_1v_1 + u_2v_2 + \cdots + u_nv_n] \\ &= [\mathbf{u}\cdot\mathbf{v}] \\ &= \mathbf{u}\cdot\mathbf{v} \end{aligned}$$

Note that we have considered the 1×1 matrix as a real number by removing its square brackets.

That is the dot product of the vectors **u** and **v** in R^n which can be expressed in the form of matrix operation as $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{u}$

Properties

If **A** is an $n \times n$ matrix, then

(i) $\mathbf{Au}\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{A}^{\mathrm{T}}\mathbf{v}$ (ii) $\mathbf{u}\cdot\mathbf{Av} = \mathbf{A}^{\mathrm{T}}\mathbf{u}\cdot\mathbf{v}$

Proof

(i) $$\begin{aligned} \mathbf{Au}\cdot\mathbf{v} &= \mathbf{v}^{\mathrm{T}}(\mathbf{Au}) = (\mathbf{v}^{\mathrm{T}}\mathbf{A})\mathbf{u} \\ &= (\mathbf{A}^{\mathrm{T}}\mathbf{v})^{\mathrm{T}}\mathbf{u} = \mathbf{u}\cdot\mathbf{A}^{\mathrm{T}}\mathbf{v} \end{aligned}$$

(ii) $$\begin{aligned} \mathbf{u}\cdot\mathbf{Av} &= (\mathbf{Av})^{\mathrm{T}}\mathbf{u} = (\mathbf{v}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}})\mathbf{u} \\ &= \mathbf{v}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{u}) \\ &= \mathbf{A}^{\mathrm{T}}\mathbf{u}\cdot\mathbf{v}. \end{aligned}$$

2. *Commutative law for vector addition:* By the definition of addition,

$$\begin{aligned}\mathbf{u}+\mathbf{v} &= [u_1+v_1, u_2+v_2, \ldots, u_n+v_n]\\ &= [v_1+u_1, v_2+u_2, \ldots, v_n+u_n] \qquad \text{(since } u_i+v_i = v_i+u_i \text{ in } R)\\ &= [v_1, v_2, \ldots, v_n]+[u_1, u_2, \ldots, u_n]\\ &= \mathbf{v}+\mathbf{u}\end{aligned}$$

i.e. $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$

3. *Associative law for vector addiition:* Let $\mathbf{w} = [w_1, w_2, \ldots, w_n]$ be any vector in R^n. Then, by the definition of addition

$$\begin{aligned}\mathbf{u}+(\mathbf{v}+\mathbf{w}) &= [u_1, u_2, \ldots, u_n]+[v_1+w_1, v_2+w_2, \ldots, v_n+w_n]\\ &= [(u_1+(v_1+w_1), u_2+(v_2+w_2), \ldots, u_n+(v_n+w_n)]\\ &= [(u_1+v_1)+w_1, (u_2+v_2)+w_1, \ldots, (u_n+v_n)+w_1)]\end{aligned}$$

Since $u_i+(v_i+w_i) = (u_i+v_i)+w_i$ in R

$$\begin{aligned} &= [u_1+v_1, u_2+v_2, \ldots, u_n+v_n]+[w_1, w_2, \ldots, w_n]\\ &= (\mathbf{u}+\mathbf{v})+\mathbf{w}\end{aligned}$$

i.e. $\mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w}$

Hence the vector addition operation is associative.

4. *Zero vector:* Take $\mathbf{0} = (0, 0, \ldots, 0)$ in R^n

$$\mathbf{u}+0 = [u_1+0, u_2+0, \ldots, u_n+0] = \mathbf{u} = \mathbf{0}+\mathbf{u}$$

i.e. $\mathbf{u}+\mathbf{0} = \mathbf{0}+\mathbf{u} = \mathbf{u}$

Therefore $\mathbf{0} = [0, 0, \ldots, 0]$ is a zero vector for the set R^n

5. *Additive inverse:* Since $u_i \in R$, so $(-u_2) \in R$

Take $-\mathbf{u} = [-u_1, -u_2, \ldots, -u_n]$ in R^n

$$\therefore \quad \begin{aligned}\mathbf{u}+(-\mathbf{u}) &= [u_1-u_1, u_2-u_2, \ldots, u_n-u_n]\\ &= [0, 0, \ldots, 0] = \mathbf{0}\end{aligned}$$

Similarly $(-\mathbf{u})+\mathbf{u} = \mathbf{0}$

i.e. $\mathbf{u}+(-\mathbf{u}) = (-\mathbf{u})+\mathbf{u} = \mathbf{0}$

Therefore each vector has its additive inverse in R^n.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication, $k\mathbf{u} = [ku_1, ku_2, \ldots, ku_n]$. Since $u_i \in R$, $i = 1, 2, 3, \ldots, n$ and k is any real scalar, so $ku_i \in R$ for each i, where $i = 1, 2, \ldots, n$. Therefore $ku_i \in R^n$. Hence R^n is closed under the given scalar multiplication.

7. *Distributive law for vector addition:* By the definition of addition,

$$\begin{aligned}k(\mathbf{u}+\mathbf{v}) &= k[u_1+v_1, u_2+v_2, \ldots, u_n+v_n]\\ &= [k(u_1+v_1), k(u_2+v_2), \ldots, k(u_n+v_n)]\\ &= [ku_1+kv_1, ku_2+kv_2, \ldots, ku_n+kv_n]\end{aligned}$$

$$= [ku_1, ku_2, \ldots, ku_n] + [kv_1, kv_2, \ldots, kv_n]$$
$$= k\mathbf{u} + k\mathbf{v}$$

i.e. $$k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$$

Therefore the addition is distributive under the scalar multiplication in R^n.

8. *Distributive law for scalar multiplication:* For any real scalars. k and m,

$$\begin{aligned}(k + m)\mathbf{u} &= (k + m)\,[u_1, u_2, \ldots, u_n]\\ &= [(k + m)u_1, (k + m)u_2, \ldots, (k + m)u_n]\\ &= [ku_1 + mu_1, ku_1 + mu_2, \ldots, ku_n + mu_n]\\ &= (ku_1, ku_2, \ldots, ku_n) + (mu_1, mu_2, \ldots, mu_n)\\ &= k[u_1, u_2, \ldots, u_n) + m(u_1, u_2, \ldots, u_n]\\ &= k\mathbf{u} + m\mathbf{u}\end{aligned}$$

Hence $$(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$$

9. *Associative law for scalar multiplication:* For any real scalars k and m,

$$\begin{aligned}k(mu) &= k(mu_1, mu_2, \ldots, mu_n)\\ &= [k(mu_1), k(mu_2), \ldots, k(mu_n)]\\ &= [(km)u_1, (km)u_2, \ldots, (km)u_n]\\ &= km[u_1, u_2, \ldots, u_n]\\ &= (km)\mathbf{u}\end{aligned}$$

Hence $$k(m\mathbf{u}) = (km)\mathbf{u}$$

10. $$\begin{aligned}1\mathbf{u} &= [1u_1, 1u_2, \ldots, 1u_n]\\ &= [u_1, u_2, \ldots, u_n]\\ &= \mathbf{u}\end{aligned}$$

i.e. $1\mathbf{u} = \mathbf{u}$

It follows that R^n is a vector space over R. This is known as the n-dimensional Euclidean space.

EXAMPLE 2.13 Let V be a set of a single vector $\mathbf{0}$ with properties $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $k\mathbf{0} = \mathbf{0}$. Then show that V is a vector space.

Solution: Here $V = \{\mathbf{0}\}$ with operations defined as follows:

$$\textit{Vector addition}:\ \mathbf{0} + \mathbf{0} = \mathbf{0}$$
$$\textit{Scalar multiplication}:\ k\mathbf{0} = \mathbf{0}$$

It is easy to show that V is a vector space …

Note: The space $V = \{\mathbf{0}\}$ is called the zero space.

EXAMPLE 2.14 Consider the set P_n of all polynomials of degree at most n with coefficients in R, i.e. $\{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mid a_0, a_1, \ldots, a_n \in R\}$. For any polynomial $\mathbf{p}(x) = p_0 + p_1x + p_2x^2 + \cdots + p_nx^n$

and $\mathbf{q}(x) = q_0 + q_1x + q_2x^2 + \cdots + q_nx^n$ in P_n and for any scalar $k \in R$, the operations on P_n are defined as below:

$$\textit{Vector addition:}\ \ \mathbf{p}(x) + \mathbf{q}(x) = (p_0 + q_0) + (p_1 + q_1)x + (p_2 + q_2)x^2 + \cdots + (p_n + q_n)x^n$$

$$\textit{Scalar multiplication:}\ \ k\mathbf{p}(x) = kp_0 + kp_1x + kp_2x^2 + \cdots + kp_nx^n$$

Show that P_n is a vector space.

Solution: A given set P_n is a vector space if it satisfies the properties of vector space.

Let $\mathbf{p}(x) = p_0 + p_1x + p_2x^2 + \cdots + p_nx^n$ and $\mathbf{q}(x) = q_0 + q_1x + q_2x^2 + \cdots + q_nx^n$ in P_n where $p_i, q_j \in R$ for each i, j with $i = 1, 2, 3, \ldots, n$, $j = 1, 2, 3, \ldots, n$ and k is any scalar.

1. *Closure property for vector addition:* By the definition of addition,

$$\mathbf{p}(x) + \mathbf{q}(x) = (p_0 + q_0) + (p_1 + q_1)x + (p_2 + q_2)x^2 + \cdots + (p_n + q_n)x^n$$

Since $p_i, q_i \in R$. So $p_i + q_i \in R$ for each $i = 1, 2, 3, \ldots, n$

Therefore, $\mathbf{p}(x) + \mathbf{q}(x) \in P_n$

2. *Commutative law for vector addition:* By the definition of addition,

$$\begin{aligned}\mathbf{p}(x) + \mathbf{q}(x) &= (p_0 + q_0) + (p_1 + q_1)x + (p_2 + q_2)x^2 + \cdots + (p_n + q_n)x^n\\ &= (q_0 + p_0) + (q_1 + p_1)x + (q_2 + p_2)x^2 + \cdots + (q_n + p_n)x^n\\ &= (q_0 + q_1x + q_2x^2 + \cdots + q_nx^n) + (p_0 + p_1x + p_2x^2 + \cdots + p_nx^n)\\ &= \mathbf{q}(x) + \mathbf{p}(x)\end{aligned}$$

$$\therefore \quad \mathbf{p}(x) + \mathbf{q}(x) = \mathbf{q}(x) + \mathbf{p}(x)$$

Hence the addition operation is commutative.

3. *Associative law for vector addition:* Let $\mathbf{r}(x) = r_0 + r_1x + r_2x^2 + \cdots + r_nx^n$ be any vector in P_n.

$$\begin{aligned}\mathbf{p}(x) + (\mathbf{q}(x) + \mathbf{r}(x)) &= (p_0 + p_1x + p_2x^2 + \cdots + p_nx^n) + ((q_0 + q_1x + q_2x^2 + \cdots + q_nx^n)\\ &\quad + (r_0 + r_1x + r_2x^2 + \cdots + r_nx^n))\\ &= ((p_0 + (q_0 + r_0)) + ((p_1 + (q_1 + r_1))x + ((p_2 + (q_1 + r_2))x^2 + \cdots\\ &\quad + ((p_n + (q_n + r_n))x^n\\ &= ((p_0 + q_0) + r_0) + ((p_1 + q_1) + r_1)x + ((p_2 + q_2) + r_2))x^2 + \cdots\\ &\quad + ((p_n + q_n) + r_n)x^n\\ &= ((p_0 + q_0) + (p_1 + q_1)x + (p_2 + q_2)x^2 + \cdots + (p_n + q_n)x^n)\\ &\quad + (r_0 + r_1x + r_2x^2 + \cdots + r_nx^n)\\ &= (\mathbf{p}(x) + \mathbf{q}(x)) + \mathbf{r}(x)\end{aligned}$$

$$\therefore \quad \mathbf{p}(x) + (\mathbf{q}(x) + \mathbf{r}(x)) = (\mathbf{p}(x) + \mathbf{q}(x)) + \mathbf{r}(x)$$

Hence the addition operation is associative.

$$= (kp_0 + kp_1x + \cdots + kp_n x^n) + (mp_0 + mp_1x + \cdots + mp_nx^n)$$
$$= k\mathbf{p}(x) + m\mathbf{p}(x)$$

Hence $\quad (k + m)\mathbf{p}(x) = k\mathbf{p}(x) + m\mathbf{p}(x)$

9. *Associative law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} k[m\mathbf{p}(x)] &= k(mp_0 + mp_1x + \cdots + mp_nx^n) \\ &= k(mp_0) + k(mp_1)x + \cdots + k(mp_nx^n) \\ &= (km)p_0 + (km)p_1x + \cdots + (km)p_nx^n \\ &= (km)\mathbf{p}(x) \end{aligned}$$

Hence $\quad k[m\mathbf{p}(x)] = (km)\mathbf{p}(x)$

10.
$$\begin{aligned} 1\mathbf{p}(x) &= 1p_0 + 1p_1x + \cdots + 1p_nx^n \\ &= p_0 + p_1x + \ldots + p_nx^n \\ &= \mathbf{p}(x) \end{aligned}$$

Hence $\quad 1\mathbf{p}(x) = \mathbf{p}(x)$

It therefore follows that P_n is a vector space over R.

EXAMPLE 2.15 Consider the set M_{22} of 2×2 matrices with operations:

$$\textit{Addition}:\ \mathbf{A} + \mathbf{B} = \begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix}$$

$$\textit{Scalar multiplication:}\ k\mathbf{A} = \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix}$$

where $\mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ and k is any real scalar. Show that M_{22} is a vector space.

Solution: Let $\mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ in M_{22} and k, m be any real scalars.

1. *Closure property for matrix addition:* By the definition of matrix addition,

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix}$$

Since $a_i, b_i \in R$, so $a_i + b_i \in R$ for each $i = 1, 2, 3, 4$

$\therefore \quad \mathbf{A} + \mathbf{B} \in M_{22}$

Hence M_{22} is closed under addition.

2. *Commutative law for matrix addition:* By the definition of matrix addition,

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} a_1+b_1 & a_2+b_2 \\ a_3+b_3 & a_4+b_4 \end{bmatrix}=\begin{bmatrix} b_1+a_1 & b_2+a_2 \\ b_3+a_3 & b_4+a_4 \end{bmatrix}=\mathbf{B}+\mathbf{A}$$

$\therefore$ $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$

Hence the addition is commutative in M_{22}.

3. *Associative law for scalar multiplication:* Let

$$\mathbf{C}=\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \text{ be any vector in } M_{22}.$$

Hence
$$\mathbf{A}+(\mathbf{B}+\mathbf{C})=\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}+\begin{bmatrix} b_1+c_1 & b_2+c_2 \\ b_3+c_3 & b_4+c_4 \end{bmatrix}$$

$$=\begin{bmatrix} a_1+(b_1+c_1) & a_2+(b_2+c_2) \\ a_3+(b_3+c_3) & a_4+(b_4+c_4) \end{bmatrix}$$

$$=\begin{bmatrix} a_1+b_1 & a_2+b_2 \\ a_3+b_3 & a_4+b_4 \end{bmatrix}+\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$$

$$=(\mathbf{A}+\mathbf{B})+\mathbf{C}$$

Thus $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$. Hence the addition operation is associative in M_{22}.

4. *Zero vector:* Consider $\mathbf{0}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ in M_{22}.

Hence
$$\mathbf{A}+\mathbf{0}=\begin{bmatrix} a_1+0 & a_2+0 \\ a_3+0 & a_4+0 \end{bmatrix}=\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}=\mathbf{A}$$

Thus $\mathbf{A}+\mathbf{0}=\mathbf{A}$.
So $\mathbf{0}$ is a zero vector for the set M_{22}.

5. *Additive inverse:* Suppose $\mathbf{B}=\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ is the additive inverse of $\mathbf{A}$ in M_{22}.

Hence
$$\mathbf{A}+\mathbf{B}=\mathbf{0}$$

or
$$\begin{bmatrix} a_1+b_1 & a_2+b_2 \\ a_3+b_3 & a_4+b_4 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

or
$$a_1+b_1=0,\quad a_2+b_2=0,\quad a_3+b_3=0 \text{ and } a_4+b_4=0$$

$\therefore$
$$b_1=-a_1,\ b_2=-a_2,\quad b_3=-a_3 \text{ and } b_4=-a_4$$

Since $a_i \in R$, so $-a_i \in R$. Hence $\mathbf{B}=-\mathbf{A}=\begin{bmatrix} -a_1 & -a_2 \\ -a_3 & -a_4 \end{bmatrix}$ is the additive inverse of $\mathbf{A}$ in M_{22}.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication,

$$kA = \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix} \text{ for any real scalar } k.$$

Since $a_i \in R$ for each i and k real scalar, so $ka_i \in R$ for each $i = 1, 2, 3, 4$.

7. *Distributive law for matrix addition:* By the definition of matrix addition,

$$\begin{aligned} k(\mathbf{A} + \mathbf{B}) &= k\begin{bmatrix} a_1 + b_1 & a_2 + b_2 \\ a_3 + b_3 & a_4 + b_4 \end{bmatrix} \\ &= \begin{bmatrix} k(a_1 + b_1) & k(a_2 + b_2) \\ k(a_3 + b_3) & k(a_4 + b_4) \end{bmatrix} \\ &= \begin{bmatrix} ka_1 + kb_1 & ka_2 + kb_2 \\ ka_3 + kb_3 & ka_4 + kb_4 \end{bmatrix} \\ &= \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix} + \begin{bmatrix} kb_1 & kb_2 \\ kb_3 & kb_4 \end{bmatrix} \\ &= k\mathbf{A} + k\mathbf{B} \end{aligned}$$

Thus $\quad k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$

8. *Distributive law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} (k + m)\mathbf{A} = (k + m)\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} &= \begin{bmatrix} (k + m)a_1 & (k + m)a_2 \\ (k + m)a_3 & (k + m)a_4 \end{bmatrix} \\ &= \begin{bmatrix} ka_1 + ma_1 & ka_2 + ma_2 \\ ka_3 + ma_3 & ka_4 + ma_4 \end{bmatrix} \\ &= \begin{bmatrix} ka_1 & ka_2 \\ ka_3 & ka_4 \end{bmatrix} + \begin{bmatrix} ma_1 & ma_2 \\ ma_3 & ma_4 \end{bmatrix} \\ &= k\mathbf{A} + m\mathbf{A} \end{aligned}$$

Hence $\quad (k + m)\mathbf{A} = k\mathbf{A} + m\mathbf{A}$

9. *Associative law for scalar multiplication:* By the defnition of scalar multiplication,

$$\begin{aligned} k(m\mathbf{A}) = k\begin{bmatrix} ma_1 & ma_2 \\ ma_3 & ma_4 \end{bmatrix} &= \begin{bmatrix} k(ma_1) & k(ma_2) \\ k(ma_3) & k(ma_4) \end{bmatrix} \\ &= \begin{bmatrix} (km)a_1 & (km)a_2 \\ (km)a_3 & (km)a_4 \end{bmatrix} = (km)\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \\ &= (km)\mathbf{A} \end{aligned}$$

Thus $k(m\mathbf{A}) = (km)\mathbf{A}$

10. $1\mathbf{A} = \begin{bmatrix} 1a_1 & 1a_2 \\ 1a_3 & 1a_4 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \mathbf{A}$

EXAMPLE 2.16 Consider the set V of real-valued functions defined on R for $f(x)$, $g(x)$ which are two functions in V and k is any real scalar. If addition and scalar multiplication are defined as

$$\textit{Addition}: \ (f+g)\,(x) = f(x) + g(x)$$
$$\textit{Scalar multiplication}: \ (kf)\,(x) = kf(x)$$

then is V a vector space with respect to the given operations?

Solution: Let $f(x)$, $g(x)$ be two functions in V and k be any real scalar.

1. *Closure property for addition:* By the definition of addition,

$$(f+g)\,(x) = f(x) + g(x)$$

Since $f(x)$, $g(x)$ are real numbers, so $f(x) + g(x)$ is also a real number. Thus, $f+g$ is a real-valued function on **R**, i.e. $f+g \in V$

2. *Commutative law for addition:* By the definition of addition,

$$\begin{aligned}(f+g)\,(x) &= f(x) + g(x)\\ &= g(x) + f(x)\\ &= (g+f)\,(x)\\ \therefore \qquad f+g &= g+f\end{aligned}$$

Thus the addition operation is commutative in V.

3. *Associative law for addition:* By the definition of addition,

$$\begin{aligned}(f+(g+h))\,(x) &= f(x) + (g+h)\,(x)\\ &= f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x)\\ &= (f+g)\,(x) + h(x)\\ &= ((f+g)+h)\,(x)\\ \therefore \qquad f+(g+h) &= (f+g)+h\end{aligned}$$

Thus the addition operation is associative in V.

4. *Zero vector:* Consider $0(x) = 0$ for all x in R.

$$\begin{aligned}(f+0)\,(x) &= f(x) + 0(x) = f(x) + 0\\ &= f(x)\\ \therefore \qquad f+0 &= f\end{aligned}$$

Thus $0(x)$ is a zero vector for V.

5. *Additive inverse:* Since $f(x) \in R$. So $-f(x) \in R$

Take
$$\begin{aligned}(-f)\,(x) &= -f(x)\\ (f+(-f))\,(x) &= f(x) + (-f)\,(x)\\ &= f(x) - f(x) = 0\\ &= 0(x)\end{aligned}$$

Thus $f+(-f) = 0$. Therefore each vector of V has its additive inverse in V.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication,

$$(kf)\,(x) = kf(x)$$

Since $f(x)$, k in R, so $kf(x)$ in R. Therefore kf is a real-valued function on R

$\therefore \qquad kf \in V$

7. *Distributive law for addition:* By the definition of addition,

$$\begin{aligned} k(f+g)(x) &= k(f(x)+g(x)) \\ &= kf(x)+kg(x) \\ &= (kf)(x)+(kg)(x) \\ &= (kf+kg)(x) \end{aligned}$$

Thus $\qquad k(f+g) = kf + kg$

8. *Distributive law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} ((k+m)f)(x) &= (k+m)f(x) \\ &= kf(x)+mf(x) \\ &= (kf+mf)(x) \end{aligned}$$

Thus $\qquad (k+m)f = kf + mf$

9. *Associative law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} k(mf)(x) &= k(mf(x)) \\ &= (km)f(x) \\ &= ((km)f)(x) \end{aligned}$$

Thus $\qquad k(mf) = (km)f$

Hence scalar multiplication in V is associative.

10. $(1f)(x) = 1f(x) = f(x)$. Thus $1f = f$

It therefore follows that the given set V of real-valued functions on **R** is a vector space.

EXAMPLE 2.17 Consider a set V of positive real numbers with the operations:

$$\textit{Addition}: \; x + y = xy$$

$$\textit{Scalar multiplication}: \; kx = x^k$$

where x, y are any two real numbers and k is any scalar.

Solution: Here $V = \{x \,|\, x > 0, x \in R\}$

Let $x, y \in V$ and k be any scalar.

The ten axioms for the given set V are as follow:

1. *Closure property for addition:* By the definition of addition,

$$x + y = xy$$

Since x, y are real numbers, so xy is also real.
Therefore $x + y \in V$.

2. *Commutative law for addition:* By the definition of addition,

$$x + y = xy = yx = y + x$$

Therefore the addition operation is commutative in V.

3. *Associative law for addition:* Let $z \in V$.

$$\begin{aligned} x + (y + z) &= x + (yz) \\ &= x(yz) \\ &= (xy)z \\ &= (x + y)z \\ &= (x + y) + z \end{aligned}$$

Therefore the addition operation is associative in V.

4. *Zero vector:* Suppose $\mathbf{0}$ is a zero vector for the vector space V. Then,

$$\begin{aligned} \mathbf{0} + \mathbf{x} &= \mathbf{x} \\ \mathbf{0x} &= \mathbf{x} = 1\mathbf{x} \\ \mathbf{0} &= [1] \end{aligned}$$

Therefore $\mathbf{0} = [1]$ is a zero vector for V.

5. *Additive inverse:* Suppose y is the additive inverse of x in V. Then,

$$\begin{aligned} y + x &= \mathbf{0} \\ yx &= 1 \\ y &= \frac{1}{x} \end{aligned}$$

Since $x \in R$, so $\frac{1}{x} \in R$. Therefore, $-x = \frac{1}{x}$ is the additive inverse of x in V.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication,

$$kx = x^k$$

Since x is a real number in V and k is also real scalar, so $x^k \in V$. That is, V is closed under scalar multiplication.

7. *Distributive law for addition:* By the definition of addition,

$$\begin{aligned} k(x + y) = k(xy) &= (xy)^k \\ &= x^k y^k \\ &= x^k + y^k \\ &= kx + ky \end{aligned}$$

8. *Distributive law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} (k + m)x &= x^{k+m} \\ &= x^k x^m \\ &= kx + mx \end{aligned}$$

Thus $$(k + m)x = kx + mx$$

9. *Associative law for scalar multiplication:* By the definition of scalar multiplication,

$$k(mx) = k(x^m)$$

$$= (x^m)^k$$
$$= x^{mk}$$
$$= (mk)\, x$$

Thus $\qquad k(mx) = (km)x$

10. $1x = x^1 = x$

It therefore follows that the given set V of real numbers is a vector space under the given operations.

EXAMPLE 2.17(a) Show that a set V of all pairs of positive real number (x, y) with operations

Addition: $(x_1, y_1) + (x_2, y_2) = (x_1x_2, y_1y_2)$

Scalar multiplication: $k(x, y) = (x^k, y^k)$, for any scalar k

is a vector space.

Solution: Here $V = \{(x, y) \mid x, y\, \varepsilon R^+\}$
Let $\mathbf{u} = [x_1, y_1]$, $\mathbf{v} = [x_2, y_2]$, $\mathbf{w} = [x_3, y_3]$ be vectors in V and k, m be any scalar.

The properties of the vector space for the set V are as follows:

1. *Closure property for addition:* By the definition of addition,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= [x_1, y_1] + [x_2, y_2] \\ &= [x_1x_2, y_1y_2] \end{aligned}$$

Since x_1, x_2, y_1, y_2 are real positive numbers, so x_1x_2 and y_1y_2 are also positive real numbers. Therefore $\mathbf{u} + \mathbf{v} \in V$. Hence V is closed under addition operation.

2. *Commutative law for addition:* By the definition of addition,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= [x_1, y_1] + [x_2, y_2] \\ &= [x_1x_2, y_1y_2] \\ &= [x_2x_1, y_2y_1] \\ &= [x_2, y_2] + [x_1, y_1] \\ &= \mathbf{v} + \mathbf{u} \end{aligned}$$

That is, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Therefore the addition operation is commutative.

3. *Associative law for addition:* By the definition of addition,

$$\begin{aligned} \mathbf{u} + (\mathbf{v} + \mathbf{w}) &= [x_1, y_1] + [x_2x_3, y_2y_3] \\ &= [x_1\,(x_2x_3), y_1(y_2y_3)] \\ &= [(x_1x_2)\,x_3, (y_1y_2)\,y_3] \\ &= [x_1x_2, y_1y_2] + [x_3, y_3] \\ &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} \end{aligned}$$

Therefore the addition is associative in V.

4. *Zero vector:* Consider $\mathbf{0} = [a, b]$ to be the zero vector for the vector space V, that is,

$$\mathbf{u} + \mathbf{0} = \mathbf{u}$$
$$[x_1, y_1] + [a, b] = [x_1, y_1]$$

$$[x_1a, y_1b] = [x_1, y_1]$$
$$x_1a = x_1, \quad y_1b = y_1$$
$$a = 1, \quad b = 1$$

Therefore $\mathbf{0} = [1, 1]$ is a zero vector for V.

5. *Additive inverse:* Suppose $\mathbf{v} = [a, b]$ is the additive inverse of $\mathbf{u} = [x_1, y_1]$. Then

$$\mathbf{u} + \mathbf{v} = \mathbf{0}$$
$$[x_1, y_1] + [a, b] = [1, 1]$$
$$[x_1a, y_1b] = [1, 1]$$
$$x_1a = 1, y_1b = 1$$
$$a = \frac{1}{x_1}, \; b = \frac{1}{y_1}$$
$$[x_1, y_1] + \left[\frac{1}{x_1}, \frac{1}{y_1}\right] = [1, 1]$$

Since x_1, y_1 in R, so $\frac{1}{x_1}, \frac{1}{y_1}$ in R. Thus $\left[\frac{1}{x_1}, \frac{1}{y_1}\right]$ is in V. Therefore the vector $\left[\frac{1}{x_1}, \frac{1}{y_1}\right]$ is the additive inverse of $[x_1, y_1]$ in V.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication,

$$k\mathbf{u} = (x^k, y^k)$$

Since x, y in R and k is also real, so kx, ky in R. Therefore, $k\mathbf{u} \in V$. Hence V is closed under scalar multiplication.

7. *Distributive law for addition:* By the definition of addition,

$$\begin{aligned} k(\mathbf{u} + \mathbf{v}) &= k[x_1x_2, y_1y_2] \\ &= [(x_1x_2)^k, (y_1y_2)^k] \\ &= [(x_1^k x_2^k, y_1^k y_2^k)] \\ &= [x_1^k, y_1^k] + [x_2^k, y_2^k] \\ &= k\mathbf{u} + k\mathbf{v} \end{aligned}$$

Thus $\quad k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

8. *Distributive law for scalar multiplication:* By the definition of scalar multiplication,

$$\begin{aligned} (k + m)\mathbf{u} = (k + m)[x_1, y_1] &= [x_1^{k+m}, y_1^{k+m}] \\ &= (x_1^k x_1^m, y_1^k y_1^m) \\ &= [x_1^k, y_1^k] + [x_1^m, y_1^m] \\ &= k\mathbf{u} + m\mathbf{u} \end{aligned}$$

Thus $\quad (k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$

9. *Associativity law for scalar multiplication:* By the definition of scalar multiplication,

$$k(m\mathbf{u}) = k[x_1^m, y_1^m] = [(x_1^m)^k, (y_1^m)^k]$$

$$= [x_1^{km}, y_1^{km}]$$
$$= (km)\,[x_1, y_1]$$
$$= (km)\mathbf{u}$$

Thus $$k(m\mathbf{u}) = (km)\mathbf{u}$$

10. $1\mathbf{u} = 1[x_1, y_1] = [x_1^1, y_1^1] = [x_1, y_1] = \mathbf{u}$

Thus $1\mathbf{u} = \mathbf{u}$

So V has satisfied all of the properties of the vector space. Hence V is a vector space under the given operations.

EXAMPLE 2.18 Show that the set W of the points on a line through the origin in R^2 with the standard addition and scalar multiplication is a vector space.

Solution: As we know, the equation of a line through the origin in R^2 is $ax + by = 0$, for some fixed real numbers a and b. Hence

$$W = \{[x, y] \in R^2 \mid ax + by = 0, \text{ for some fixed real numbers } a \text{ and } b\}$$

The axioms of the vector space are as follows.

Let $\mathbf{u} = [x_1, y_1]$, $\mathbf{v} = [x_2, y_2] \in W$ and k be any scalar. That is, $ax_1 + by_1 = 0$, $ax_2 + by_2 = 0$

1. *Closure:* By the definition of addition in R^2, we have

$$\mathbf{u} + \mathbf{v} = [x_1 + x_2,\ y_1 + y_2]$$
$$a(x_1 + x_2) + b(y_1 + y_2) = (ax_1 + by_1) + (ax_2 + by_2)$$
$$= 0 + 0 \qquad (\text{since } \mathbf{u}, \mathbf{v} \in W)$$
$$= \mathbf{0}$$

Thus $\mathbf{u} + \mathbf{v} \in W$. Therefore W is closed under addition operation.

2. *Commutative property:* Since W is a subset of R^2 and R^2 is commutative with respect to addition operation, so W is also commutative with respect to addition.

3. *Associative property:* Since R^2 is associative under addition, so the subset W of R^2 is associative.

4. *Zero vector:* Take $\mathbf{0} = [0, 0]$ in R^2.

$$ax + by = a(0) + b(0)$$
$$= \mathbf{0}$$

Therefore $\mathbf{0} = [0, 0]$ is a vector of W and also

$$\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}.$$

Hence $\mathbf{0} = [0, 0]$ is a zero vector for W.

5. *Additive inverse:* Suppose $\mathbf{v} = [x_2, y_2]$ is an additive inverse of $\mathbf{u} = [x_1, y_1]$, then

$$\mathbf{u} + \mathbf{v} = \mathbf{0}$$
$$[x_1 + x_2, y_1 + y_2] = [0, 0]$$
$$x_2 = -x_1,\ y_2 = -y_1$$

and also

$$a(-x_1) + b(-y_1) = -ax_1 - by_1$$

$$= -(ax_1 + by_1)$$
$$= 0 \qquad (\text{since } ax_1 + by_1 = 0 \text{ as } \mathbf{u} = [x_1, y_1] \text{ is in } W)$$

Thus $\mathbf{v} = -\mathbf{u} = [-x_1, -y_1]$ is a vector of W. So $\mathbf{v} = -\mathbf{u} = [-x_1, -y_1]$ is an additive inverse of $\mathbf{u} = [x_1, y_1]$ in W.

6. *Closure property for scalar multiplication:* By the definition of scalar multiplication,

$$k\mathbf{u} = [kx_1, ky_1]$$
$$a(kx_1) + b(ky_1) = k(ax_{1+}\text{by}_1)$$
$$= k(0) \qquad (\text{since } ax_1 + by_1 = 0 \text{ as } \mathbf{u} = [x_1, y_1] \text{ in } W)$$

Thus $k\mathbf{u} \in W$. Hence W is closed under scalar multiplication.

7. *Distributive property for addition:* By the definition of addition,

$$k(\mathbf{u} + \mathbf{v}) = [kx_1 + kx_2, ky_1 + ky_2]$$
$$a(kx_1 + kx_2) + b(ky_1 + ky_2) = k(ax_1 + by_1) + k(ax_2 + by_2)$$
$$= 0$$

Thus $k(\mathbf{u} + \mathbf{v}) \in W$. Since addition is distributive in R^2, so it is in W.

8. *Distributive property for scalar addition:* Let m be any real scalar.

$$(k + m)\mathbf{u} = [(k + m)x_1, (k + m)y_1]$$
$$a(k + m)x_1 + b(k + m)y_1 = (k + m)(ax_1 + by_1)$$
$$= (k + m)(0) \qquad (\text{since } ax_1 + by_1 = 0)$$
$$= 0$$

Thus $(k + m)\mathbf{u} \in W$. Since scalar multiplication is distributive in R^2, so it is in W.

9. *Associative property for scalar multiplication:* By the definition of scalar multiplication,

$$k(m\mathbf{u}) = [k(mx_1), k(my_1)]$$
$$= [kmx_1, kmy_1]$$
$$a(kmx_1) + b(kmy_1) = km(ax_1 + by_1)$$
$$= km(0) \qquad (\text{since } ax_1 + by_1 = 0)$$
$$= 0$$

Thus $k(m\mathbf{u})$ is a vector of W. Since scalar multiplication is associative in R^2, so it is in W.

10. $1\mathbf{u} = (1x_1, 1y_1) = (x_1, y_1) \in W$

It follows from the above axioms that a set W of the points on a line through the origin in R^2 with the standard addition and scalar multiplication is a vector space.

EXAMPLE 2.19 Is the set V of all pairs of real numbers (x, y) with the operations

Addition: $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2 + 1, y_1 + y_2 + 1)$

Scalar multiplication: $k(x, y) = (kx, ky)$

a vector space? If yes, then prove it; if no, then identify all axioms that fail to hold.

Solution: The given set V is not a vector space under the given operations because it does not satisfy the axiom 7, i.e. the distributive property:

Let $\mathbf{u} = [x_1, y_1],\ \mathbf{v} = [x_2, y_2]$

$$k(\mathbf{u} + \mathbf{v}) = k[x_1 + x_2 + 1, y_1 + y_2 + 1]$$
$$= [kx_1 + kx_2 + k, ky_1 + ky_2 + k]$$
$$k\mathbf{u} + k\mathbf{v} = [kx_1, ky_1] + [kx_2, ky_2]$$
$$= [kx_1 + kx_2 + 1, ky_1 + ky_2 + 1]$$

$\therefore$ $$k(\mathbf{u} + \mathbf{v}) \neq k\mathbf{u} + k\mathrm{v}.$$

In the next discussion, we shall try to find the answers to the following questions:

(i) Is one vector space contained in another vector space?
(ii) Can we find a vector space which contains a given set of vectors?

The answer to first question is in the following definition.

Definition: Subspace

Let W be a non-empty subset of a vector space V. Then W is said to be subspace of V if W is a vector space under the same operations of addition and scalar multiplication as those in V.

EXAMPLE 2.20 Is the set of all points in R^2 lying on a line which passes through the origin, a subspace of R^2?

Solution: Take $V = R^2$

and W = the set of all points in R^2 lying on a line which passes through the origin.

In Example 2.18, we showed that W is a vector space under the same operations of coordinate-wise addition and scalar multiplication as in R^2.

Therefore by the definition, W is a subspace of R^2. By the definition of subspace, a set W is a subspace of a vector space V; we need to check all of the properties of the vector space. But some of the properties like commutative, associative, distributive are inherited from V. So, we need not verify those properties. Hence, we have the following theorem.

Theorem 2.7 [Subspace]

Let W be a non-empty subset of a vector space V. Then W is a subspace of V if and only if W satisfies the following properties:

(i) If $\mathbf{u}, \mathbf{v} \in W$, then $\mathbf{u} + \mathbf{v} \in W$
(ii) If $\mathbf{u} \in W$ and k is any scalar, then $k\mathbf{u} \in W$

EXAMPLE 2.21 Show that the set $W = \{(1, x) | x \in R\}$ is a subspace of R^2 under the operations

Addition: $[1, x] + [1, y] = [1, x + y]$

Scalar multiplication: $k[1, x] = [1, kx]$, k is any scalar

Solution: By Theorem 2.7, it is sufficient to show that W is closed under the given operation of addition and scalar multiplication.

Let $\mathbf{u} = [1, x],\ \mathbf{v} = [1, y]$ be two vectors in W.

$$\mathbf{u} + \mathbf{v} = [1, x] + [1, y] = [1, x + y] \in W$$
$$k\mathbf{u} = k[1, x] = [1, kx] \in W$$

Thus W is closed under addition and scalar multiplication, so it is a subspace of R^2.

EXAMPLE 2.22 Show that the set $W = \{[x, y, z] \ | y = x + z\}$ is a subspace of R^3 under the usual addition and scalar multiplication.

Solution: Let $\mathbf{u} = [x_1, y_1, z_1]$, $\mathbf{v} = [x_2, y_2, z_2]$ be two vectors in W.

$$\mathbf{u} + \mathbf{v} = [x_1, y_1, z_1] + [x_2, y_2, z_2]$$
$$= [x_1 + x_2,\ \ y_1 + y_2,\ \ z_1 + z_2]$$

Since $\mathbf{u}$, $\mathbf{v}$ are the vectors of W. Therefore,

$$y_1 = x_1 + z_1,\ \ y_2 = x_2 + z_2$$
$$y_1 + y_2 = (x_1 + z_1) + (x_2 + z_2)$$
$$y_1 + y_2 = (x_1 + x_2) + (z_1 + z_2)$$

Thus $\mathbf{u} + \mathbf{v} \in W$

$$k\mathbf{u} = k[x_1, y_1, z_1] = [kx_1, ky_1, kz_1]$$
$$ky_1 = k(x_1 + z_1)$$
$$ky_1 = kx_1 + kz_1$$

Thus $k\mathbf{u} \in W$.
Therefore W is closed under addition and scalar multiplication, so it is a subspace of R^3.

EXAMPLE 2.23 Is the set $W = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \middle| \, a_{11} + a_{12} + a_{21} + a_{22} = 0 \right\}$ a subspace of M_{22} under the usual addition and scalar multiplication?

Solution: Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ be two vectors of W.

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

To prove that $\mathbf{A} + \mathbf{B}$ is a vector of W, it is sufficient to show that the sum of its elements is zero.

$$(a_{11} + b_{11}) + (a_{12} + b_{12}) + (a_{21} + b_{21}) + (a_{22} + b_{22})$$
$$= (a_{11} + a_{12} + a_{21} + a_{22}) + (b_{11} + b_{12} + b_{21} + b_{22})$$
$$= 0 + 0 \qquad \text{(since } \mathbf{A} \text{ and } \mathbf{B} \text{ are the vectors of } W\text{)}$$
$$= 0$$

Thus $\mathbf{A} + \mathbf{B} \in W$.

Let k be any scalar.

$$k\mathbf{A} = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$$

$$ka_{11} + ka_{12} + ka_{21} + ka_{22} = k(a_{11} + a_{12} + a_{21} + a_{22})$$
$$= k(0) \qquad \text{(since } \mathbf{A} \text{ is a vector of } W\text{)}$$
$$= 0$$

Thus $k\mathbf{A} \in W$. Hence a subset W of M_{22} is a subspace of M_{22} under the usual addition and scalar multiplication.

EXAMPLE 2.24 Is a set $W = \{[\mathbf{A}]_{nxn} | \text{ tr } \mathbf{A} = 0\}$ a subspace of M_{nn} under the usual addition and scalar multiplication?

Solution: Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$ be two vectors of W.

Therefore,

$$\text{tr } \mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn} = 0$$

$$\text{tr } \mathbf{B} = b_{11} + b_{22} + \cdots + b_{nn} = 0$$

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nn}+b_{nn} \end{bmatrix}$$

$$\begin{aligned} \text{tr } (\mathbf{A} + \mathbf{B}) &= (a_{11} + b_{11}) + (a_{22} + b_{22}) + \cdots + (a_{nn} + b_{nn}) \\ &= (a_{11} + a_{22} + \cdots + a_{nn}) + (b_{11} + b_{22} + \cdots + b_{nn}) \\ &= 0 + 0 \qquad (\text{since } \mathbf{A}, \mathbf{B} \in W) \\ &= 0 \end{aligned}$$

$\therefore$ $\mathbf{A} + \mathbf{B} \in W$.

Let k be any scalar. Then,

$$k\mathbf{A} = k \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ ka_{n1} & ka_{n2} & \cdots & ka_{nn} \end{bmatrix}$$

$$\begin{aligned} \text{tr } (k\mathbf{A}) &= ka_{11} + ka_{22} + \cdots + ka_{nn} \\ &= k(a_{11} + a_{22} + \cdots + a_{nn}) \\ &= k(0) \qquad \text{since } \mathbf{A} \in W \\ &= 0 \end{aligned}$$

$\therefore$ $k\mathbf{A} \in W$.

Thus W is closed under addition and scalar multiplication, so it is a subspace of M_{nn}.

EXAMPLE 2.25 Consider a subset $W = \{p(x) \in P_n \mid p(0) + p'(0) = 0\}$ of P_n with the usual addition and scalar multiplication of polynomials. Show that W is a subspace of P_n.

Solution: Let $\mathbf{p}(x)$, $\mathbf{q}(x)$ be two vectors in W. Therefore,

$$p(0) + p'(0) = 0 \quad \text{and} \quad q(0) + q'(0) = 0.$$

By the definition of addition in P_n,

$$(p + q)(x) = p(x) + q(x)$$

Moreover
$$(p+q)'(x) = p'(x) + q'(x)$$
$$\begin{aligned}(p+q)(0) + (p+q)'(0) &= (p(0)+q(0)) + (p'(0)+q'(0))\\ &= (p(0)+p'(0)) + (q(0)+q'(0))\\ &= 0+0 \qquad (\text{since } p(x), q(x) \in W)\\ &= 0\end{aligned}$$

Therefore $(p+q)(x)$ is a vector in W.
Let k be any scalar
By the definition of scalar multiplication in P_n,
$$(kp)(x) = kp(x)$$
Moreover
$$(kp)'(x) = kp'(x)$$
$$\begin{aligned}(kp)(0) + (kp)'(0) &= kp(0) + kp'(0)\\ &= k(p(0)+p'(0))\\ &= k(0) \qquad (\text{since } p(x) \in W)\\ &= 0\end{aligned}$$

Therefore $(kp)(x)$ is a vector in W.

Thus W is closed under addition and scalar multiplication, so it is a subspace of P_n.

EXAMPLE 2.26 Consider the set $C(-\infty, \infty)$ of all continuous functions on the internal $(-\infty, \infty)$ with operations,
$$\textit{Addition: } (f+g)(x) = f(x) + g(x)$$
$$\textit{Scalar multiplication: } (kf)(x) = kf(x)$$
where f, g are vectors of $C(-\infty, \infty)$ and k is any scalar. Show that a set W of all constant functions is a subspace of $C(-\infty, \infty)$.

Solution: Let f, g be two constant functions on $(-\infty, \infty)$. Therefore,
$$f(x) = c_1,\ g(x) = c_2$$
where c_1 and c_2 are two real constants.
By the definition of addition,
$$\begin{aligned}(f+g)(x) &= f(x) + g(x)\\ &= c_1 + c_2\\ &= \text{constant}\end{aligned}$$
Therefore, $f+g$ is a vector of W.
By the definition of scalar multiplication,
$$\begin{aligned}(kf)(x) &= kf(x)\\ &= kc_1\\ &= \text{constant}\end{aligned}$$
Therefore, kf is a vector of W.

Thus, W is closed under addition and scalar multiplication, so it is a subspace of $C(-\infty, \infty)$.

EXAMPLE 2.27 Show that any plane through the origin in R^3 is a subspace of R^3.

Solution: A plane through the origin can be represented as

$$W = \{[x, y, z] \mid ax + by + cz = 0\ ,\ a, b, c \in R\}$$

Let $\mathbf{u} = [x_1, y_1, z_1]$, $\mathbf{v} = [x_2, y_2, z_2]$ be two vectors in W

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= [x_1, y_1, z_1] + [x_2, y_2, z_2] \\ &= (x_1 + x_2, y_1 + y_2,\ z_1 + z_2)\end{aligned}$$

and also

$$\begin{aligned}&a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) \\ &= (ax_1 + by_1 + cz_1) + (ax_2 + by_2 + cz_2) \\ &= 0 + 0 \qquad (\text{since } \mathbf{u}, \mathbf{v} \in W) \\ &= 0\end{aligned}$$

Therefore $\mathbf{u} + \mathbf{v}$ is a vector of W.

$$k\mathbf{u} = k[x_1, y_1, z_1] = [kx_1, ky_1, kz_1]$$

and also

$$\begin{aligned}&a(kx_1) + b(ky_1) + c(kz_1) \\ &= k(ax_1 + by_1 + cz_1) \\ &= k(0) \qquad (\text{since } \mathbf{u} \in W)\end{aligned}$$

Therefore $k\mathbf{u}$ is a vector of W.

Thus W is closed under addition and scalar multiplication, so that W is a subspace of R^3.

EXAMPLE 2.28 Is the set $W = \{[a, 1, 1] \mid a \in R\}$subspace of R^3 under the usual addition and scalar multiplication?

Solution: Let $\mathbf{u} = [a, 1, 1]$, $\mathbf{v} = [b, 1, 1]$ be two vectors of R^3.

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= [a, 1, 1] + [b, 1, 1] \\ &= [a + b, 2, 2]\end{aligned}$$

$\mathbf{u} + \mathbf{v}$ is not of the form $(a, 1, 1)$. So, $\mathbf{u} + \mathbf{v}$ is not a vector in W. Therefore W is not a subspace of R^3.

EXAMPLE 2.29 Is the set $W = \{\mathbf{A} \in M_{22} \mid \det \mathbf{A} = 0\}$ a subspace of M_{22} under the usual addition and scalar multiplication?

Solution: In general, $\det(\mathbf{A} + \mathbf{B}) \neq \det \mathbf{A} + \det \mathbf{B}$

For example, $\mathbf{A} = \begin{bmatrix} 4 & 10 \\ 2 & 5 \end{bmatrix}$ with $\det \mathbf{A} = 0$ and $\mathbf{B} = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$ with $\det \mathbf{B} = 0$

But $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 6 & 13 \\ 8 & 14 \end{bmatrix}$ with $\det(\mathbf{A} + \mathbf{B}) = -20 \neq 0$

Therefore $\mathbf{A}$ and $\mathbf{B}$ are the vectors of W but $\mathbf{A} + \mathbf{B}$ is not a vector of W. So, W is not a subspace of M_{22}.

Linear Combination

Let us start with the definition.

Definition: Linear Combination

If a vector $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$

then the vector $\mathbf{w}$ is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars.

EXAMPLE 2.30 Express a vector $\mathbf{w} = [7, 4, -3]$ as a linear combination of $\mathbf{v}_1 = [1, -2, -5]$ and $\mathbf{v}_2 = [2, 5, 6]$.

Solution: Suppose for some scalar α_1, α_2,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{w}$$
$$\alpha_1[1, -2, -5] + \alpha_2 [2, 5, 6] = [7, 4, -3]$$

Equating the corresponding components gives

$$\begin{aligned} \alpha_1 + 2\alpha_2 &= 7 \\ -2\alpha_1 + 5\alpha_2 &= 4 \\ -5\alpha_1 + 6\alpha_2 &= -3 \end{aligned}$$

We solve this system of equations by the Gauss-elimination method.

$$\begin{bmatrix} 1 & 2 & : & 7 \\ -2 & 5 & : & 4 \\ -5 & 6 & : & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & : & 7 \\ 0 & 9 & : & 18 \\ 0 & 16 & : & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & : & 7 \\ 0 & 1 & : & 2 \\ 0 & 0 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & : & 3 \\ 0 & 1 & : & 2 \\ 0 & 0 & : & 0 \end{bmatrix}$$

Thus the solutions are $\alpha_1 = 3$, $\alpha_2 = 2$.

$\therefore$ $3[1, -2, -5] + 2[2, 5, 6] = [7, 4, -3]$, i.e. $3\mathbf{v}_1 + 2\mathbf{v}_2 = \mathbf{w}$

EXAMPLE 2.31 Which of the following vectors are the linear combinations of $\mathbf{v}_1 = [0, -2, 2]$ and $\mathbf{v}_2 = [1, 3, -1]$?

(i) [2, 2, 2] (ii) [0, 4, 5]

Solution:

(i) If $\mathbf{w}_1 = [2, 2, 2]$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, then there exists α_1 and α_2 such that $\mathbf{w}_1 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$

or $$[2, 2, 2] = \alpha_1[0, -2, 2] + \alpha_2[1, 3, -1]$$

Equating the corresponding components gives

$$\begin{aligned} \alpha_2 &= 2 \\ -2\alpha_1 + 3\alpha_2 &= 2 \\ 2\alpha_1 - \alpha_2 &= 2 \end{aligned}$$

We solve this system by the Gauss-elimination method

$$\begin{bmatrix} 0 & 1 & : & 2 \\ -2 & 3 & : & 2 \\ 2 & -1 & : & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & : & 2 \\ -2 & 0 & : & -4 \\ 2 & 0 & : & 4 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & : & 2 \\ 1 & 0 & : & 2 \\ 0 & 0 & : & 0 \end{bmatrix}$$

Thus the solutions are $\alpha_1 = 2$, $\alpha_2 = 2$. So,

$\therefore$ $$\mathbf{w}_1 = 2\mathbf{v}_1 + 2\mathbf{v}_2$$

i.e. $$[2, 2, 2] = 2[0, -2, 2] + 2[1, 3, -1]$$

(ii) If $\mathbf{w}_2 = [0, 4, 5]$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, then there exist α_1 and α_2 such that

$$\mathbf{w}_2 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$$

or $$[0, 4, 5] = \alpha_1[0, -2, 2] + \alpha_2[1, 3, -1]$$

Equating the corresponding components gives

$$\begin{aligned} \alpha_2 &= 0 \\ -2\alpha_1 + 3\alpha_2 &= 4 \\ 2\alpha_1 - \alpha_2 &= 5 \end{aligned}$$

We solve this system by the Gauss-elimination method

$$\begin{bmatrix} 0 & 1 & : & 0 \\ -2 & 3 & : & 4 \\ 2 & -1 & : & 5 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & : & 0 \\ -2 & 0 & : & 4 \\ 2 & 0 & : & 5 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & : & 1 \\ 2 & 0 & : & -4 \\ 0 & 0 & : & 9 \end{bmatrix}$$

Thus the system is inconsistent, so such scalars α_1 and α_2 do not exist.

EXAMPLE 2.32 Express $\mathbf{q} = -9 - 7x - 15x^2$ as a linear combination of $\mathbf{p}_1 = 2 + x + 4x^2$, $\mathbf{p}_2 = 1 - x + 3x^2$ and $\mathbf{p}_3 = 3 + 2x + 5x^2$.

Solution: In order for $\mathbf{q} = -9 - 7x - 15x^2$ to be a linear combination of $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$, there must be scalars α_1, α_2 and α_3 such that

$$\mathbf{q} = \alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3$$

or $$-9 - 7x - 15x^2 = \alpha_1(2 + x + 4x^2) + \alpha_2(1 - x + 3x^2) + \alpha_3(3 + 2x + 5x^2)$$

or $$-9 - 7x - 15x^2 = (2\alpha_1 + \alpha_2 + 3\alpha_3) + (\alpha_1 - \alpha_2 + 2\alpha_3)x + (4\alpha_1 + 3\alpha_2 + 5\alpha_3)x^2$$

Equating the corresponding components gives

$$\begin{aligned} 2\alpha_1 + \alpha_2 + 3\alpha_3 &= -9 \\ \alpha_1 - \alpha_2 + 2\alpha_3 &= -7 \\ 4\alpha_1 + 3\alpha_2 + 5\alpha_3 &= -15 \end{aligned}$$

Using the Gauss-elimination method, we have

$$\begin{bmatrix} 2 & 1 & 3 & : & -9 \\ 1 & -1 & 2 & : & -7 \\ 4 & 3 & 5 & : & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 & : & -7 \\ 2 & 1 & 3 & : & -9 \\ 4 & 3 & 5 & : & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 & : & -7 \\ 0 & 3 & -1 & : & 5 \\ 0 & 1 & -1 & : & 3 \end{bmatrix} \sim$$

$$\left[\begin{array}{rrrcr} 1 & -1 & 2 & : & -7 \\ 0 & 1 & -1 & : & 3 \\ 0 & 0 & 2 & : & -4 \end{array}\right] \sim \left[\begin{array}{rrrcr} 1 & 0 & 0 & : & -2 \\ 0 & 1 & 0 & : & 1 \\ 0 & 0 & 1 & : & -2 \end{array}\right]$$

Thus the solutions are $\alpha_1 = -2$, $\alpha_2 = 1$, $\alpha_3 = -2$. Hence

$$\mathbf{q} = -2\mathbf{p}_1 + \mathbf{p}_2 - 2\mathbf{p}_3$$

i.e. $\quad -9 - 7x - 15x^2 = -2\,(2 + x + 4x^2) + 1(1 - x + 3x^2) - 2(3 + 2x + 5x^2)$

EXAMPLE 2.33 Is $\mathbf{U} = \begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ a linear combination of $\mathbf{A} = \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} -12 & 2 \\ 1 & 4 \end{bmatrix}$?

Solution: Suppose **U** is linear combination of **A**, **B** and **C**. Then there exist scalars α_1, α_2 and α_3 such that

$$\mathbf{U} = \alpha_1\mathbf{A} + \alpha_2\mathbf{B} + \alpha_3\mathbf{C}$$

or

$$\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix} = \alpha_1 \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + \alpha_3 \begin{bmatrix} -12 & 2 \\ 1 & 4 \end{bmatrix}$$

By comparing the corresponding components, we get

$$\begin{aligned} 4\alpha_1 + \alpha_2 - 12\alpha_3 &= 6 \\ -\alpha_2 + 2\alpha_3 &= 0 \\ -2\alpha_1 + 2\alpha_2 + \alpha_3 &= 3 \\ -2\alpha_1 + 3\alpha_2 + 4\alpha_3 &= 8 \end{aligned}$$

Using the Gauss-elimination method, we have

$$\left[\begin{array}{rrrcr} 4 & 1 & -12 & : & 6 \\ 0 & -1 & 2 & : & 0 \\ -2 & 2 & 1 & : & 3 \\ -2 & 3 & 4 & : & 8 \end{array}\right] \sim \left[\begin{array}{rrrcr} 0 & 5 & -10 & : & 12 \\ 0 & -1 & 2 & : & 0 \\ 0 & -1 & -3 & : & -5 \\ -2 & 3 & 4 & : & 8 \end{array}\right] \sim \left[\begin{array}{rrrcr} 0 & 0 & 0 & : & 12 \\ 0 & -1 & 2 & : & 0 \\ 0 & -1 & -3 & : & -5 \\ -2 & 3 & 4 & : & 8 \end{array}\right]$$

The first row of the row-reduced form suggests that this system is inconsistent, so no such α_1, α_2 and α_3 scalars exist.

Span

Definition: Span

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a subset of a vector space V. Then the set of all linear combinations of the vectors in S is called the *span* of S. It is denoted by span (S) and it can also be represented as

$$\begin{aligned} \text{span } (S) &= \text{span } \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\} \\ &= \{\mathbf{w} \mid \mathbf{w} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_r\mathbf{v}_r, \text{ each } \alpha_i \text{ is scalar, } 1 \le i \le r\} \end{aligned}$$

EXAMPLE 2.34 Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two vectors of a vector space V. Then show that $W = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ is a subspace of V.

Solution: To prove W is a subspace of V, we have to show the following results:

(i) The zero vector is in W.
(ii) W is closed under vector addition.
(iii) W is closed under scalar multiplication.

The zero vector is in W, since $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$.
If $\mathbf{w}_1$ and $\mathbf{w}_2$ are in W, then

$$\mathbf{w}_1 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \quad \text{and} \quad \mathbf{w}_2 = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 \text{ for some scalars } \alpha_1, \alpha_2, \beta_1, \beta_2.$$

$$\begin{aligned}\mathbf{w}_1 + \mathbf{w}_2 &= (\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) + (\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2)\\ &= (\alpha_1 + \beta_1)\mathbf{v}_1 + (\alpha_2 + \beta_2)\mathbf{v}_2\end{aligned}$$

So $\mathbf{w}_1 + \mathbf{w}_2$ is in W.
For any scalar k,

$$k\mathbf{w}_1 = k(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = (k\alpha_1)\mathbf{v}_1 + (k\alpha_2)\mathbf{v}_2$$

which shows that $k\mathbf{w}_1$ is in W. Thus W is closed under addition and scalar multiplication, and so W is a subspace of a vector space V.

If we extend Example 2.34 to more than two vectors, then we have the following theorem.

Theorem 2.8 [Span of Vectors]

If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are vectors in a vector space V, then

(i) span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is subspace of V.
(ii) span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is the smallest subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, that is, every other subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ must contain span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$.

EXAMPLE 2.35 Which of the following sets of vectors span R^3.

(i) $S = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$
(ii) $S_1 = \{[1, 1, 1], [0, 1, 1], [0, 0, 1], [1, 2, 3]\}$

Solution:

(i) Let $\mathbf{v}_1 = [1, 0, 0]$, $\mathbf{v}_2 = [0, 1, 0]$ and $\mathbf{v}_3 = [0, 0, 1]$
If a set S spans R^3, then an arbitrary vector $\mathbf{b} = [b_1, b_2, b_3]$ of R^3 can be expressed as a linear combination,

$$\mathbf{b} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 \quad \text{for some scalars } \alpha_1, \alpha_2 \text{ and } \alpha_3.$$

$$[b_1, b_2, b_3] = \alpha_1[1, 0, 0] + \alpha_2[0, 1, 0] + \alpha_3[0, 0, 1]$$

$$[b_1, b_2, b_3] = [\alpha_1, \alpha_2, \alpha_3]$$

By comparing the corresponding components, we get

$$\alpha_1 = b_1, \alpha_2 = b_2, \alpha_3 = b_3$$

$$[b_1, b_2, b_3] = b_1[1, 0, 0] + b_2[0, 1, 0] + b_3[0, 0, 1]$$

Thus an arbitrary vector $\mathbf{b}$ of R^3 can be expressed as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$. So the set S spans R^3, that is, $R^3 = \text{span}(S)$.

(ii) Let $\mathbf{w}_1 = [1, 1, 1]$, $\mathbf{w}_2 = [0, 1, 1]$, $\mathbf{w}_3 = [0, 0, 1]$ and $\mathbf{w}_4 = [1, 2, 3]$.

If a set S_1 spans $\mathbf{R}^3$, then an arbitrary vector $\mathbf{b} = [b_1, b_2, b_3]$ of $\mathbf{R}^3$ can be expressed as a linear combination, $\mathbf{b} = \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \beta_3\mathbf{w}_3 + \beta_4\mathbf{w}_4$

For some real scalars β_1, β_2, β_3 and β_4,

$$[b_1, b_2, b_3] = \beta_1[1, 1, 1] + \beta_2[0, 1, 1] + \beta_3[0, 0, 1] + \beta_4[1, 2, 3]$$

Equating the corresponding components, we have

$$\begin{aligned} \beta_1 + \beta_4 &= b_1 \\ \beta_1 + \beta_2 + 2\beta_4 &= b_2 \\ \beta_1 + \beta_2 + \beta_3 + 3\beta_4 &= b_3 \end{aligned}$$

By applying the Gauss-elimination method to this system, we have

$$\begin{bmatrix} 1 & 0 & 0 & 1 & : & b_1 \\ 1 & 1 & 0 & 2 & : & b_2 \\ 1 & 1 & 1 & 3 & : & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & : & b_1 \\ 1 & 1 & 0 & 2 & : & b_2 \\ 0 & 0 & 1 & 1 & : & b_3 - b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & : & b_1 \\ 0 & 1 & 0 & 1 & : & b_2 - b_1 \\ 0 & 0 & 1 & 1 & : & b_3 - b_2 \end{bmatrix}$$

$$\begin{aligned} \beta_1 + \beta_4 &= b_1 \\ \beta_2 + \beta_4 &= b_2 - b_1 \\ \beta_3 + \beta_4 &= b_3 - b_2 \end{aligned}$$

If we choose $\beta_4 = 0$, then $\beta_1 = b_1$, $\beta_2 = b_2 - b_1$ and $\beta_3 = b_3 - b_2$. Thus,

$$[b_1, b_2, b_3] = b_1[1, 1, 1] + (b_2 - b_1)\,[0, 1, 1] + (b_3 - b_2)\,[0, 0, 1] + 0[1, 2, 3]$$

$$\mathbf{b} = b_1\mathbf{w}_1 + (b_2 - b_1)\mathbf{w}_2 + (b_3 - b_2)\mathbf{w}_3$$

Thus an arbitrary vector $\mathbf{b}$ of R^3 can be expressed as a linear combination of vectors of a set S_1. So the set S_1 spans R^3, that is, $R^3 = \text{span}\,(S_1)$.

In this example, we have seen that R^3 has two different spanning sets. In general, spanning sets are not unique. One obvious question arises in the mind. Is there any relation between the spanning sets of a vector space V? The answer to this question is given in the following theorem.

Theorem 2.9 [Relationship between the Spanning Sets of a Vector Space]

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $S_1 = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are two sets of vectors in a space V, then span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ = span $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ if and only if each vector in S is a linear combination of those in S_1 and each vector in S_1 is a linear combination of those in S.

EXAMPLE 2.36 Do $\mathbf{p}_1 = 2 + 2x + 2x^2$, $\mathbf{p}_2 = 3x^2$ and $\mathbf{p}_3 = x + x^2$ span P_2?

Solution: Suppose $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ span P_2. Then an arbitrary vector $\mathbf{q} = b_1 + b_2x + b_3x^2$ can be expressed as a linear combination,

$$\mathbf{q} = \alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3$$

For some scalars α_1, α_2 and α_3,

$$b_1 + b_2x + b_3x^2 = \alpha_1(2 + 2x + 2x^2) + \alpha_2(3x^2) + \alpha_3(x + x^2)$$

By comparing the coefficients of powers of x, we have

$$2\alpha_1 = b_1$$

$$2\alpha_1 + \alpha_3 = b_2$$

$$2\alpha_1 + 3\alpha_2 + \alpha_3 = b_3$$

Using the Gauss-elimination method,

$$\begin{bmatrix} 2 & 0 & 0 & : & b_1 \\ 2 & 0 & 1 & : & b_2 \\ 2 & 3 & 1 & : & b_3 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 & : & b_1 \\ 2 & 0 & 1 & : & b_2 \\ 0 & 3 & 0 & : & b_3 - b_2 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 & : & b_1 \\ 0 & 0 & 1 & : & b_2 - b_1 \\ 0 & 1 & 0 & : & \dfrac{b_3 - b_2}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & : & \dfrac{b_1}{2} \\ 0 & 0 & 1 & : & b_2 - b_1 \\ 0 & 1 & 0 & : & \dfrac{b_3 - b_2}{3} \end{bmatrix}$$

Thus, $$\alpha_1 = \frac{b_1}{2}, \alpha_3 = b_2 - b_1, \alpha_3 = \frac{b_3 - b_2}{3}$$

$\therefore$ $$\mathbf{q} = \frac{b_1}{2}\mathbf{p}_1 + (b_2 - b_1)\mathbf{p}_2 + \frac{b_3 - b_2}{3}\mathbf{p}_3$$

Thus an arbitrary polynomial of P_2 can be expressed as a linear combination of $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$. So span (P_2) = span $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

EXAMPLE 2.37 Suppose $f = \cos^2 x$ and $g = \sin^2 x$, then which of the following lie in the space spanned by f and g? (i) $\cos 2x$ (ii) 1.

Solution:

(i) Since $\cos 2x = \cos^2 x - \sin^2 x$, we have

$$\cos 2x = f - g$$

Thus for $\alpha_1 = 1$ and $\alpha_2 = -1$, $\cos 2x$ is expressed as a linear combination of $f = \cos^2 x$ and $g = \sin^2 x$. So, $\cos 2x$ lies in the space spanned by f and g.

(ii) Since $1 = \cos^2 x + \sin^2 x$, for $\alpha_1 = 1$ and $\alpha_2 = 1$, we have 1 expressed as a linear combination of f and g. So 1 lies in the space spanned by f and g.

EXAMPLE 2.38 For what values of h will $\mathbf{w} = [-4, 3, h]$ be in the subspace of R^3 spanned by $\mathbf{v}_1 = [1, -1, 2]$, $\mathbf{v}_2 = [5, -4, 13]$ and $\mathbf{v}_3 = [-3, 1, -12]$.

Solution: For some scalars $\alpha_1, \alpha_2, \alpha_3$

$$\mathbf{w} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3$$

or $$[-4, 3, h] = \alpha_1[1, -1, 2] + \alpha_2[5, -4, 13] + \alpha_3[-3, 1, -12]$$

or $$[-4, 3, h] = [\alpha_1 + 5\alpha_2 - 3\alpha_3, -\alpha_1 - 4\alpha_2 + \alpha_3, 2\alpha_1 + 13\alpha_2 - 12\alpha_3]$$

By comparing the corresponding components,

$$\begin{aligned} \alpha_1 + 5\alpha_2 - 3\alpha_3 &= -4 \\ -\alpha_1 - 4\alpha_2 + \alpha_3 &= 3 \\ 2\alpha_1 + 13\alpha_2 - 12\alpha_3 &= h \end{aligned}$$

Using the Gauss-elimination method,

$$\begin{bmatrix} 1 & 5 & -3 & : & -4 \\ -1 & -4 & 1 & : & 3 \\ 2 & 13 & -12 & : & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & : & -4 \\ 0 & 1 & -2 & : & -1 \\ 0 & 3 & -6 & : & h+8 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & : & -4 \\ 0 & 1 & -2 & : & -1 \\ 0 & 0 & 0 & : & h+11 \end{bmatrix}$$

Since the system is consistent. So $h + 11 = 0$, i.e. $h = -11$.

EXERCISE SET 2

1. Determine which sets are vector spaces under the given operations:

(i) The set of all triples of real numbers $[x, y, z]$ with the operations

$$\textit{Addition:}\ [x, y, z] + [x', y', z'] = [x + x', y + y', z + z']$$
$$\textit{Scalar multiplication:}\ k[x, y, z] = [0, 0, 0].$$

(ii) The set of all 2×2 matrices of the form $\mathbf{A} = \begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with standard matrix operations of addition and scalar multiplication.

(iii) The set of all real numbers $[x, y]$ with operations: $[x, y] + [x', y'] = [xx', yy']$ and $k[x, y] = [kx, ky]$.

2. Show that the set of all points in R^3 lying on a plane is a vector space, with standard operations of addition and scalar multiplication, in which the plane passes through the origin.

3. The set of all 2×2 matrices of the form $\begin{bmatrix} a & a+1 \\ b+1 & b \end{bmatrix}$ with standard matrix operations of addition and scalar multiplication, is a vector space . Verify it.

4. Show that the set of vectors of different dimensions is not a vector space.

5. Show that the set of polynomials of the form $a + bx + cx^2 + dx^3$ with the standard polynomial operations of addition and scalar multiplication, is a vector space.

6. Show that a set $W = \{[x, y, z] | z = x + y\}$ is a subspace of R^3 under the usual addition and scalar multiplication operations.

7. Determine which of the following are the subspaces of R^3.

(i) All the vectors of the form $[0, 0, c]$

(ii) All the vectors of the form $c = a + b + 1$.

(iii) All the vectors of the form $[a, 0, c]$

8. Determine whether matrices $\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$ of the form are subspaces of M_{22} with the usual addition and scalar multiplication operations.

9. Determine which of the following is the subspace of the space $F(-\infty, \infty)$

(i) All the constant functions

(ii) All the functions such that $f(x) \le 0$, for all x

(iii) All the functions such that $f(0) = a$, where a is a constant.

(iv) All f and g such that $f(0) = g(0) = 1$

10. Determine which of the following can be expressed as linear combination of $[-1, 2, 3]$ and $[2, -4, 6]$

(i) $[-3, 6, 0]$ (ii) $[2, 4, 18]$ (iii) $[0, 0, 0]$

11. Show that $[2, 1, 4]$ and $[1, -1, 2]$ are a linear combination of $[3, 1, -2]$.

12. Show that $[1, 2, 1, 0]$, $[1, 3, 1, 2]$ and $[4, 2, 1, 0]$ are linear combination of $[6, 1, 0, 1]$.

13. Show that $[1, 1, 0]$ and $[0, 1, 1]$ are a linear combination $[1, 0, 1]$.

14. Determine whether $[1, 1, 2]$, $[1, 0, 1]$ and $[2, 1, 3]$ span R^3.

15. Which of the following sets of vectors span R^3?

(i) $\{[1, 0, 2], [1, -1, -1], [3, 2, 4]\}$

(ii) $\{[1, 1, 0], [0, 1, 1], [1, 0, 1]\}$

(iii) $\{[2, 1, -1], [1, -7, -8], [1, 3, 2]\}$.

2.3 LINEAR INDEPENDENCE AND DEPENDENCE OF VECTORS

In the previous section, we discussed two concepts: linear combination and span. We also learned that a set of vector spaces gives a vector space V if every vector of V can be expressed as a linear combination of vectors of S, but this expression may not be unique. For example, a set

$$S = \{[1, 0], [0, 1], [1, 2]\}$$

spans a vector space R^2.

If we take a vector [2, 3] of R^2, then

$$[2, 3] = 2[1, 0] + 3[0, 1] - 0[1, 2]$$
$$[2, 3] = 3[1, 0] + 5[0, 1] - 1[1, 2]$$
$$[2, 3] = 4[1, 0] + 7[0, 1] - 2[1, 2]$$
$$[2, 3] = 5[1, 0] + 9[0, 1] - 3[1, 2]$$

That is, a vector can be expressed as a linear combination of vectors of S in more than one way. Can it be expressed uniquely? In this section we will try to find the answer to this question.

Let us start with some definitions.

Definition: Linearly Independent Vectors

A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of vectors is said to be *linearly independent*, if the vector equation:

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$$

has only one trivial solution $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. If this vector equation has other non-trivial solutions, that is, $\alpha_1, \alpha_2, \ldots, \alpha_n$ not all zero, then S is called a *linearly dependent* set.

EXAMPLE 2.39 Which of the following sets of vectors in R^2 are linearly independent?

(i) $S_1 = \{[1, 0], [0, 1]\}$

(ii) $S_2 = \{[1, 0], [0, 1], [a, b]\}$

(iii) $S_3 = \{[1, 3], [0, 5], [1, 5]\}$

(iv) $S_4 = \{[1, 3], [2, 5], [0, 0]\}$

Solution:

(i) Here $S_1 = \{[1, 0], [0, 1]\}$. Say, $\mathbf{v}_1 = [1, 0]$, $\mathbf{v}_2 = [0, 1]$.

Consider the equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0} \qquad \text{where } \alpha_1, \alpha_2, \text{ are scalars.}$$

The vector equation in component form is

$$\alpha_1[1, 0] + \alpha_2[0, 1] = [0, 0]$$
$$[\alpha_1, \alpha_2] = [0, 0]$$
$$\therefore \qquad \alpha_1 = 0, \alpha_2 = 0$$

Therefore, the vector equation has only the trivial solution, i.e. $\alpha_1 = \alpha_2 = 0$.

Hence the given set S_1 is a linearly independent set.

(ii) Here S_2 ={[1, 0], [0, 1], [a, b[}. Say, $\mathbf{v}_1$ = [1, 0], $\mathbf{v}_2$ = [0, 1], $\mathbf{v}_3$ = [a, b].
Consider the vector equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{where } \alpha_1, \alpha_2 \text{ and } \alpha_3 \text{ are scalars.}$$

The vector equation in component form is

$$\alpha_1[1, 0] + \alpha_2[0, 1] + \alpha_3[a, b] = [0, 0]$$
$$(\alpha_1 + \alpha_3 a, \alpha_2 + \alpha_3 b) = (0, 0)$$
$$\therefore \qquad \alpha_1 + \alpha_3 a = 0, \quad \alpha_2 + \alpha_3 b = 0$$

If we fixed $\alpha_3 = 1$, then $\alpha_1 = -a$, $\alpha_2 = -b$. That is fixed.

$$(-a)\mathbf{v}_1 + (-b)\mathbf{v}_2 + 1\mathbf{v}_3 = \mathbf{0}$$
$$(-a)\,[1, 0] + (-b)\,[0, 1] + [a, b] = [0, 0]$$

Therefore the vector equation has a non-trivial solution, so $\alpha_1 = -a$, $\alpha_2 = -b$, and $\alpha_3 = 1$. Hence the given set S_2 is a linearly dependent set.

(iii) S_3 = {[1, 3], [0.5, 1.5]}
Let $\mathbf{v}_1$= [1, 3], $\mathbf{v}_2$ = [0.5, 1.5]

By the method of inspection, we can write

$$[1, 3] = 2[0.5, 1.5]$$
$$1\mathbf{v}_1 + (-2)\mathbf{v}_2 = \mathbf{0}$$

Therefore the vector equation has no trivial solution, so S_3 is a linearly dependent set.

(iv) S_4 = {[1, 3], [2, 3], [0, 0]}
Let $\mathbf{v}_1$ = [1, 3], $\mathbf{v}_2$ = [2, 3] and $\mathbf{v}_3$ = [0, 0].

The given vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ can be expressed as

$$0[1, 3] + 0[2, 3] + 1[0, 0] = [0, 0]$$
$$0\,\mathbf{v}_1 + 0\,\mathbf{v}_2 + 1\,\mathbf{v}_3 = \mathbf{0}$$

Therefore the vector equation has the non-trivial solution $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 1$. So, S_4 is a linearly dependent set.

EXAMPLE 2.40 Which of the following sets of vectors in R^3 are linearly independent?

(i) S_1 = {[1, 0, 0], [0, 1, 0], [0, 0, 1]}
(ii) S_2 = {[1, 0, 0], [0, 1, 0], [0, 0, 1], [a, b, c]}
(iii) S_3 = {[1, 2, 5], [3, 6, 5]}
(iv) S_4 = {[1, 0, 1], [0, 1, 0], [0, 0, 0]}

Solution:

(i) Let $\mathbf{v}_1$ = [1, 0, 1], $\mathbf{v}_2$ = [0, 1, 0] and $\mathbf{v}_3$ = [0, 0, 1].

For some scalars α_1, α_2 and α_3, the vector equation is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

The component form of the vector equation is

$$\alpha_1[1, 0, 0] + \alpha_2[0, 1, 0] + \alpha_3[0, 0, 1] = [0, 0, 0]$$

$$[\alpha_1, \alpha_2, \alpha_3] = [0, 0, 0]$$

$$\therefore \qquad \alpha_1 = \alpha_2 = \alpha_3 = 0$$

Therefore the vector equation has only the trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$.
So the given set S_1 is linearly independent.

(ii) Here $S_2 = \{[1, 0, 0], [0, 1, 0], [0, 0, 1], [a, b, c]\}$
Let $\mathbf{v}_1 = [1, 0, 0]$, $\mathbf{v}_2 = [0, 1, 0]$, $\mathbf{v}_3 = [0, 0, 1]$, $\mathbf{v}_4 = [a, b, c]$
For same scalars, α_1, α_2, α_3 and α_4 the vector equation is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = \mathbf{0}$$

The component form of the vector equation is

$$\alpha_1[1, 0, 0] + \alpha_2[0, 1, 0] + \alpha_2[0, 0, 1] + \alpha_4[a, b, c] = [0, 0, 0]$$

$$(\alpha_1 + \alpha_4 a,\ \ \alpha_2 + \alpha_4 b,\ \ \alpha_3 + \alpha_4 c) = (0, 0, 0)$$

$$\therefore \qquad \alpha_1 + \alpha_4 a = 0,\ \ \alpha_2 + \alpha_4 b = 0,\ \ \alpha_3 + \alpha_4 c = 0,$$

If we fixed $\alpha_4 = 1$, $\alpha_1 = -a$, $\alpha_2 = -b$, $\alpha_3 = -c$,

i.e. $(-a)\,[1, 0, 0] + (-b)\,[0, 1, 0] + (-c)\,[0, 0, 1] + [a, b, c] = [0, 0, 0]$

or $[a, b, c] = a[1, 0, 0] + b[0, 1, 0] + c[0, 0, 1]$

Therefore the vector equation has the non-trivial solution,

$$\alpha_1 = -a,\ \alpha_2 = -b,\ \alpha_3 = -c \text{ and } \alpha_4 = 1$$

So the given set S_2 is linearly dependent.

Moreover vector $\mathbf{v}_4$ is expressed as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$.

(iii) Here $S_3 = \{[1, 2, 5], [3, 6, 15]\}$
Let $\mathbf{v}_1 = [1, 2, 5]$, $\mathbf{v}_2 = [3, 6, 15]$.
By the method of inspection, we can write

$$(3, 6, 15) = 3(1, 2, 5)$$

$$\therefore \qquad 1[3, 6, 15] - 3[1, 2, 5] = [0, 0, 0]$$

$$\text{or} \qquad (-3)\mathbf{v}_1 + 1\mathbf{v}_2 = \mathbf{0}$$

Therefore the vector equation has the non-trivial solution. So S_3 is a linear dependent set

(iv) Here $S_4 = \{[1, 0, 1], [0, 1, 0], [0, 0, 0]\}$
Let $\mathbf{v}_1 = [1, 0, 1]$, $\mathbf{v}_2 = [0, 1, 0]$, $\mathbf{v}_3 = [0, 0, 0]$.
The given vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ can be expressed as

$$0[1, 0, 1] + 0[0, 1, 0] + 1[0, 0, 0] = [0, 0, 0]$$

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + 1\mathbf{v}_3 = \mathbf{0}$$

Therefore the vector equation has the non-trivial solution, $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 1$. So, S_4 is a linearly dependent set.

If we observe the above two examples minutely, then we can come out with some results which are stated in the following theorem.

Theorem 2.10 [Linear Dependence and Linear Independance of Vectors]

A set S with two or more vector is

(i) linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.

(ii) linearly independent if and only if no vector in S is expressible as linear combination of the other vectors in S.

Corollary 1 A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.

Corollary 2 A finite set of vectors that contains the zero vector is linearly dependent.

Recall

(i) A homogeneous system $\mathbf{Ax} = \mathbf{0}$ of n linear equations in n unknowns has the trivial solution if $\det \mathbf{A} \neq 0$

(ii) A homogeneous system $\mathbf{Ax} = \mathbf{0}$ of n linear equations in m unknowns has the non-trivial solution if $m > n$.

EXAMPLE 2.41 Which of the following sets of vectors of $R^3$are independent?

(i) $S_1 = \{[-3, 0, 4], [5, -1, 2], [1, 1, 3]\}$

(ii) $S_2 = \{[1, 2, 3], [3, 2, 1], [3, 3, 3]\}$

(iii) $S_3 = \{[-2, 0, 1], [3, 2, 5], [6, -1, 1], [7, 0, -2]\}$

Solution:

(i) Consider the given set

$$S_1 = \{[-3, 0, 4], [5, -1, 2], [1, 1, 3]\}$$

Let $\mathbf{v}_1 = [-3, 0, 4]$, $\mathbf{v}_2 = [5, -1, 2]$ and $\mathbf{v}_3 = [1, 1, 3]$.
The vector equation for the vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

for some scalars $\alpha_1, \alpha_2, \alpha_3$.
The component form of the vector equation is

$$\alpha_1[-3, 0, 4] + \alpha_2[5, -1, 2] + \alpha_3[1, 1, 3] = [0, 0, 0]$$

By comparing the corresponding components, we get

$$\begin{aligned} -3\alpha_1 + 5\alpha_2 + \alpha_3 &= 0 \\ -\alpha_2 + \alpha_3 &= 0 \\ 4\alpha_1 + 2\alpha_2 + 3\alpha_3 &= 0 \end{aligned}$$

The matrix form of the above homogeneous system of linear equations is

$$\mathbf{Ax} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} -3 & 5 & 1 \\ 0 & -1 & 1 \\ 4 & 2 & 3 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} -3 & 5 & 1 \\ 0 & -1 & 1 \\ 4 & 2 & 3 \end{vmatrix} = 39 \neq 0$$

Therefore det $\mathbf{A} \neq 0$, so the system has only the trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Hence S_1 is a linearly independent set.

(ii) Consider the given set

$$S_2 = \{[1, 2, 3], [3, 2, 1], [3, 3, 3]\}$$

Let $\mathbf{v}_1 = [1, 2, 3]$, $\mathbf{v}_2 = [3, 2, 1]$ and $\mathbf{v}_3 = [3, 3, 3]$.
The vector equation for the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

The component form of the vector equation is

$$\alpha_1[1, 2, 3] + \alpha_2[3, 2, 1] + \alpha_3[3, 3, 3] = [0, 0, 0]$$

By comparing the corresponding components, we get

$$\begin{aligned} \alpha_1 + 3\alpha_2 + 3\alpha_3 &= 0 \\ 2\alpha_1 + 2\alpha_2 + 3\alpha_3 &= 0 \\ 3\alpha_1 + \alpha_2 + 3\alpha_3 &= 0 \end{aligned}$$

The matrix form of above homogeneous system of linear equations is

$$\mathbf{Ax} = \mathbf{0}$$

where $\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 3 & 1 & 3 \end{bmatrix}$, $\quad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$

$$\det \mathbf{A} = \begin{vmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 3 & 1 & 3 \end{vmatrix} = 0$$

Since det $\mathbf{A} = 0$, the system has the non-trivial solution, that is, the vector equation has the non-trivial solution. Hence S_2 is a linearly dependent set.

(iii) Consider the given set

$$S_3 = \{[-2, 0, 1], [3, 2, 5], [6, -1, 1], [7, 0, -2]\}$$

Let $\mathbf{v}_1 = [-2, 0, 1]$, $\mathbf{v}_2 = [3, 2, 5]$, $\mathbf{v}_3 = [6, -1, 1]$, $\mathbf{v}_4 = [7, 0, -2]$.
The vector equation for the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = \mathbf{0}$$

for some scalars α_1, α_2, α_3, α_4.
The components form of the vector equation is

$$\alpha_1[-2, 0, 1] + \alpha_2[3, 2, 5] + \alpha_3[6, -1, 1] + \alpha_4[7, 0, -2] = [0, 0, 0]$$

By comparing the corresponding components, we have

$$\begin{aligned} -2\alpha_1 + 3\alpha_2 + 6\alpha_3 + 7\alpha_4 &= 0 \\ 2\alpha_2 - \alpha_3 \qquad &= 0 \\ \alpha_1 + 5\alpha_2 + \alpha_3 - 2\alpha_4 &= 0 \end{aligned}$$

The homogeneous linear equation has four unknowns α_1, α_2, α_3, α_4 and three equations. So it has at least one non-trivial solution. Therefore the vector equation has also a non-trivial solution. Hence the S_3 is a linearly dependent set.

EXAMPLE 2.42 Which of the following sets of vectors are independent?

(i) $S_1 = \{[1, 0, 0, 0], [2, 5, 0, 0], [-1, -2, 1, 0], [1, 0, 4, 0]\}$

(ii) $S_2 = \{[1, 3, 1, 2], [0, 1, 0, 0], [0, 2, -2, 0], [-1, 2, 1, 1]\}$

(iii) $S_3 = \{[0, 0, 0, 0, 3], [0, 0, 0, 3, 0], (0, 0, 2, 1, 4], [0, -6, 5, 0, 1], [-1, 0, 6, 3, 1]\}$

Solution:

(i) Consider the given set

$$S_1 = \{[1, 0, 0, 0], [2, 5, 0, 0], [-1, -2, 1, 0], [1, 0, 4, 0]\}$$

Let $\mathbf{v}_1 = [1, 0, 0, 0]$, $\mathbf{v}_2 = [2, 5, 0, 0]$, $\mathbf{v}_3 = [-1, -2, 1, 0]$, $\mathbf{v}_4 = [1, 0, 4, 0]$.
The vector equation of vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = \mathbf{0}$$

for some scalars α_1, α_2, α_3, α_4.
In the component form,

$$\alpha_1[1, 0, 0, 0] + \alpha_2[2, 5, 0, 0] + \alpha_3[-1, -2, 1, 0] + \alpha_4[1, 0, 4, 0] = [0, 0, 0, 0]$$

By comparing the corresponding components,

$$\begin{aligned} \alpha_1 + 2\alpha_2 - \alpha_3 + \alpha_4 &= 0 \\ 5\alpha_2 - 2\alpha_3 &= 0 \\ \alpha_3 + 4\alpha_4 &= 0 \end{aligned} \qquad \text{(i)}$$

The matrix form of the system (i) is

$$\mathbf{A}\mathbf{x} = 0$$

$$\text{where } \mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 5 & -2 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the matrix **A** is an upper triangular matrix, so, the det **A** is the value of multiplication entries. Hence

$$\det \mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 5 & -2 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= (1)\,(5)\,(1)\,(0)$$

$$= 0$$

Since det $\mathbf{A} = 0$, the system (i) has the non-trivial solution. Therefore, the vector equation has also the non-trivial solution. Hence the given set S_1 is a linearly dependent set.

(ii) Consider the given set

$$S_2 = \{[1, 3, 1, 2], [0, 1, 0, 0], [0, 2, -2, 0], [-1, 2, 1, 1]\}$$

Let $\mathbf{v}_1 = [1, 3, 1, 2]$, $\mathbf{v}_2 = [0, 1, 0, 0]$, $\mathbf{v}_3 = [0, 2, -2, 0]$ and $\mathbf{v}_4 = [-1, 2, 2, 1]$.

The vector equation of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = \mathbf{0}$$

for some scalars α_1, α_2, α_3, α_4. Hence

$$\alpha_1[1, 3, 1, 2] + \alpha_2[0, 1, 0, 0] + \alpha_3[0, 2, -2, 0] + \alpha_4[-1, 2, 1, 1] = [0, 0, 0, 0]$$

By comparing the corresponding components, we have

$$\begin{aligned} \alpha_1 \qquad\qquad\qquad - \alpha_4 &= 0 \\ 3\alpha_1 + \alpha_2 + 2\alpha_3 + 2\alpha_4 &= 0 \\ \alpha_1 \qquad - 2\alpha_3 + \alpha_4 &= 0 \\ 2\alpha_1 \qquad\qquad\qquad + \alpha_4 &= 0 \end{aligned} \tag{i}$$

The vector form of the system (i) is

$$\mathbf{A}\mathbf{x} = 0$$

where $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 3 & 1 & 2 & 2 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & 0 & 1 \end{bmatrix}$

Here we will use the cofactors method to find det $\mathbf{A}$. Hence

$$\det \mathbf{A} = \begin{vmatrix} 1 & 0 & 0 & -1 \\ 3 & 1 & 2 & 2 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & 0 & 1 \end{vmatrix}$$

$$= 1\begin{vmatrix} 1 & 2 & 2 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{vmatrix} - 0 + 0 - (-1)\begin{vmatrix} 3 & 1 & 2 \\ 1 & 0 & -2 \\ 2 & 0 & 0 \end{vmatrix}$$

$$= 1(-2) + (-1)(0 + 4)$$

$$= -6 \neq 0$$

Since det $\mathbf{A} \neq 0$, the system has only the trivial solution, that is, the vector equation has only the trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Hence S_2 is a linearly independent set.

(iii) Consider the given set

$$S_3 = \{[0, 0, 0, 0, 3], [0, 0, 0, 3, 0], [0, 0, 2, 1, 4], [0, -6, 5, 0, 1], [-1, 0, 6, 3, 1]\}.$$

Let $\mathbf{v}_1 = [-1, 0, 6, 3, 1]$, $\mathbf{v}_2 = [0, -6, 5, 0, 1]$, $\mathbf{v}_3 = [0, 0, 2, 1, 4]$, $\mathbf{v}_4 = [0, 0, 0, 3, 0]$, $\mathbf{v}_5 = [0, 0, 0, 0, 3]$.

The vector equation of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ and $\mathbf{v}_5$ is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 + \alpha_5\mathbf{v}_5 = \mathbf{0}$$

for some scalars α_1, α_2, α_3, α_4 and α_5.
The component form of the vector equation is

$$\begin{aligned}&\alpha_1[-1, 0, 6, 3, 1] + \alpha_2[0, -6, 5, 0, 1] + \alpha_3[0, 0, 2, 1, 4]\\ &+ \alpha_4[0, 0, 0, 3, 0] + \alpha_5[0, 0, 0, 0, 3] = [0, 0, 0, 0, 0]\end{aligned}$$

By comparing the components, we have

$$\begin{aligned}-\alpha_1 \qquad\qquad\qquad\qquad &= 0\\ -6\alpha_2 \qquad\qquad\qquad &= 0\\ 6\alpha_1 + 5\alpha_2 + 2\alpha_3 \qquad\qquad &= 0\\ 3\alpha_1 \qquad + \alpha_3 + 3\alpha_4 \qquad &= 0\\ \alpha_1 + \alpha_2 + 4\alpha_3 \qquad + 3\alpha_5 &= 0\end{aligned} \tag{i}$$

The matrix form of the system (i) is

$$\mathbf{Ax} = 0$$

$$\text{where } \mathbf{A} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0\\ 0 & -6 & 0 & 0 & 0\\ 6 & 5 & 2 & 0 & 0\\ 3 & 0 & 1 & 3 & 0\\ 1 & 1 & 4 & 0 & 3\end{bmatrix}; \qquad \mathbf{x} = \begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_5\end{bmatrix}$$

Since the matrix **A** is a lower triangular matrix, hence

$$\begin{aligned}\det \mathbf{A} &= \begin{vmatrix} -1 & 0 & 0 & 0 & 0\\ 0 & -6 & 0 & 0 & 0\\ 6 & 5 & 2 & 0 & 0\\ 3 & 0 & 1 & 3 & 0\\ 1 & 1 & 4 & 0 & 3\end{vmatrix}\\ &= (-1)(-6)(2)(3)(3)\\ &= 108\\ &\neq 0\end{aligned}$$

Since $\det \mathbf{A} \neq 0$, the system has only the trivial solution, that is, the vector equation has only the trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$. Hence the set S_3 is a linearly independent set.

EXAMPLE 2.43 Which of the following sets of vectors in P_2 are linearly independent?

(i) $S_1 = \{2 - x + 4x^2, 3 + 6x + 2x^2, 2 + 10x - 4x^2\}$
(ii) $S_2 = \{1 + 3x + 3x^2, x + 4x^2, 5 + 6x + 3x^2, 7 + 2x - x^2\}$

Solution:

(i) $S_1 = \{2 - x + 4x^2, 3 + 6x + 2x^2, 2 + 10x - 4x^2\}$
Let $\mathbf{p}_1 = 2 - x + 4x^2$, $\mathbf{p}_2 = 3 + 6x + 2x^2$, $\mathbf{p}_3 = 2 + 10x - 4x^2$.

The vector equation of vectors $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ is

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 = \mathbf{0}$$

$$\alpha_1(2 - x + 4x^2) + \alpha_2(3 + 6x + 2x^2) + \alpha_3(2 + 10x - 4x^2) = 0 + 0x + 0x^2$$

By comparing the coefficients of powers of x, we have

$$\begin{aligned} 2\alpha_1 + 3\alpha_2 + 2\alpha_3 &= 0 \\ -\alpha_1 + 6\alpha_2 + 10\alpha_3 &= 0 \\ 4\alpha_1 + 2\alpha_2 - 4\alpha_3 &= 0 \end{aligned} \qquad \text{(i)}$$

The matrix form of system (i) is

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 2 \\ -1 & 6 & 10 \\ 4 & 2 & -4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 2 & 3 & 2 \\ -1 & 6 & 10 \\ 4 & 2 & -4 \end{vmatrix} = -32 \neq 0$$

Since det $\mathbf{A} \neq 0$, the system (i) has only the trivial solution, that is, the vector equation has only the trivial solution. Hence the set S_1 is a linearly independent set.

(ii) Here $S_2 = \{1 + 3x + 3x^2, x + 4x^2, 5 + 6x + 3x^2, 7 + 2x - x^2\}$

Let $\mathbf{p}_1 = 1 + 3x + 3x^2$, $\mathbf{p}_2 = x + 4x^2$, $\mathbf{p}_3 = 5 + 6x + 3x^2$, and $\mathbf{p}_4 = 7 + 2x - x^2$.

The vector equation of vectors $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ is

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 + \alpha_4\mathbf{p}_4 = \mathbf{0}$$

$$\alpha_1(1 + 3x + 3x^2) + \alpha_2(x + 4x^2) + \alpha_3(5 + 6x + 3x^2) + \alpha_4(7 + 2x - x^2) = 0 + 0x + 0x^2$$

By comparing the coefficients of powers of x, we have

$$\begin{aligned} a_1 \qquad + 5a_3 + 7a_4 &= 0 \\ 3a_1 + a_2 + 6a_3 + 2a_4 &= 0 \\ 3a_1 + 4a_2 + 3a_3 - a_4 &= 0 \end{aligned} \qquad \text{(i)}$$

The homogeneous linear system (i) has three equations in four unknowns, so it has at least only the non-trivial solution, that is, the vector equation has the non-trivial solution. Hence the set S_2 is a linearly dependent set.

If we closely observe the above examples, then we come out with the following result.

Theorem 2.11 [Linear Dependence]

If the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of vectors of R^n (or P_n) and $m > n$, then S is linearly dependent.

Geometrical Meaning of Linear Independence

Linear Independence in R^2

Consider two vectors $\mathbf{v}_1 = [a_{11}, a_{21}]$ and $\mathbf{v}_2 = [a_{12}, a_{22}]$ in R^2. To study linear independence, we consider the vector equation

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0}$$

which can be written in matrix form as

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent precisely, when

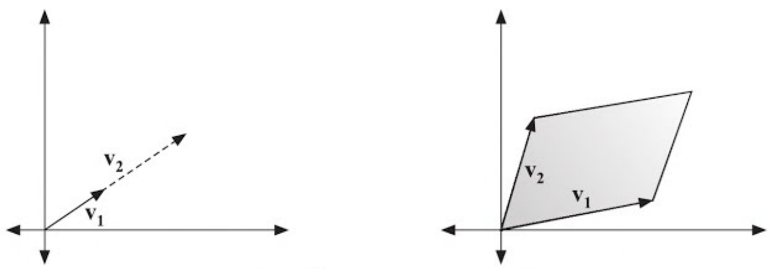
$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) \neq 0$$

This can be interpreted geometrically in the following way:
The area of the parallelogram formed by the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is in fact equal to the absolute value of the determinant of the matrix formed with $\mathbf{v}_1$ and $\mathbf{v}_2$ as the columns. In other words, if

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = 0$$

it means that the two vectors are linearly dependent precisely. That is, when the parallelogram has zero area and hence the two vectors lie on the same line [Figure 2.6(a)].

On the other hand, if $\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) \neq 0,$ the parallelogram has a positive area, that is, the two vectors are linearly independent (Figure 2.6(b)).



(a) Linearly dependent vectors in R^2 (b) Linearly independent vectors in R^2

Figure 2.6 Geometrical interpretation of vectors in R^2.

Linear Independence in R^3

Consider three vectors $\mathbf{v}_1 = [a_{11}, a_{21}, a_{31}]$, $\mathbf{v}_2 = [a_{12}, a_{22}, a_{32}]$ and $\mathbf{v}_3 = [a_{13}, a_{23}, a_{33}]$ in R^3. To study linear independence, we consider the vector equation

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

which can be written in matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent precisely, when

$$\det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) \neq 0$$

This can be interpreted geometrically in the following way:

The volume of the parallelepiped formed by the three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ is in fact equal to the absolute value of the determinant of the matrix formed with $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ as the columns. In other words, when the

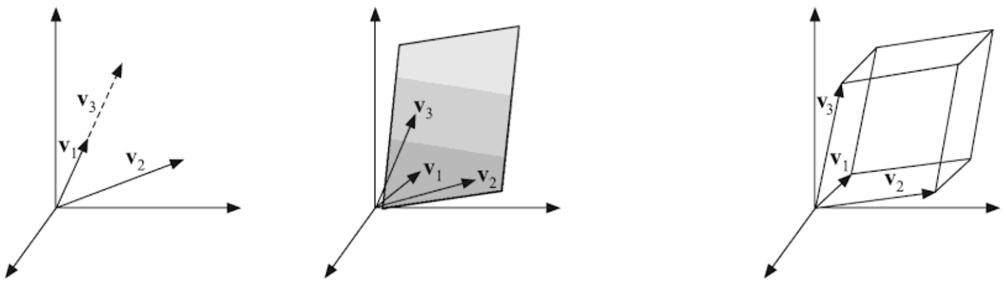
$$\det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) = 0$$

it means that the three vectors are linearly dependent precisely, when the parallelepiped has zero volume; that is, when the three vectors lie on the same plane (Figure 2.7(a)).

On the other hand, if

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \neq 0,$$

the parallelepiped has a positive volume, i.e. the three vectors are linearly independent [Figure 2.7(b)].



(a) Linearly dependent vectors in R^3. (b) Linearly independent vectors in R^3.

Figure 2.7 Geometrical interpretation of vectors in R^3.

Linear Independence of Two Vectors in R^3

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are nonzero and linearly dependent, then α_1, $\alpha_2 \in R$ exist, not both zero, such that $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0}$. This forces the two vectors to be multiples of each other, so that they lie on the same line, hence the parallelogram they form has zero area. It follows that if two vectors in R^3 form a parallelogram with positive area, then they are linearly independent.

EXAMPLE 2.44 Determine whether the three vectors of R^3 lie in a plane. (Here assume that they have their initial points at the origin.)

(i) $\mathbf{v}_1 = [2, -2, 0]$, $\mathbf{v}_2 = [6, 1, 4]$, $\mathbf{v}_3 = [2, 0, 4]$
(ii) $\mathbf{v}_1 = [-6, 7, 2]$, $\mathbf{v}_2 = [3, 2, 4]$, $\mathbf{v}_3 = [4, -1, 2]$

Solution: From the geometrical meaning of linearly independence, the three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ in R^3 lie on the same plane only when they are linearly dependent vectors. In other words,

$$\det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) = 0$$

where $\mathbf{v}_1 = [a_{11}, a_{21}, a_{31}]$, $\mathbf{v}_2 = [a_{12}, a_{22}, a_{32}]$ and $\mathbf{v}_3 = [a_{13}, a_{23}, a_{33}]$ in R^3.

(i) $\mathbf{v}_1 = [2, -2, 0]$, $\mathbf{v}_2 = [6, 1, 4]$, $\mathbf{v}_3 = [2, 0, 4]$

Hence
$$\det\begin{bmatrix} 2 & 6 & 2 \\ -2 & 1 & 0 \\ 0 & 4 & 4 \end{bmatrix} = 40 \neq 0$$

Therefore the vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent, i.e. they cannot lie on the same plane.

(ii) For $\mathbf{v}_1 = [-6, 7, 2]$, $\mathbf{v}_2 = [3, 2, 4]$, $\mathbf{v}_3 = [4, -1, 2]$, we have

$$\det\begin{bmatrix} -6 & 3 & 4 \\ 7 & 2 & -1 \\ 2 & 4 & 2 \end{bmatrix} = 0$$

Therefore the vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly dependent, i.e. they do lie on a plane.

EXAMPLE 2.45 Determine whether the three vectors of R^3 lie on the same line. (Here assume that they have their initial points at the origin.)

(i) $\mathbf{v}_1 = [3, -4, 1]$, $\mathbf{v}_2 = [-3, 4, -1]$, $\mathbf{v}_3 = [6, -8, 2]$
(ii) $\mathbf{v}_1 = [3, 2, -1]$, $\mathbf{v}_2 = [-6, -4, 2]$, $\mathbf{v}_3 = [0, 6, -3]$
(iii) $\mathbf{v}_1 = [1, 4, 1]$, $\mathbf{v}_2 = [3, 1, -1]$, $\mathbf{v}_3 = [4, 5, 0]$

Solution: The three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ of R^3 do lie on the same line if $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ can be expressed as multiple of each other, and the values of α_1, α_2 and α_3 of the vector equation are dependent on each other.

(i) For $\mathbf{v}_1 = [3, -4, 1]$, $\mathbf{v}_2 = [-3, 4, -1]$, $\mathbf{v}_3 = [6, -8, 2]$, the vector equation is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2 \text{ and } \alpha_3$$

$$\alpha_1[3, -4, 1] + \alpha_2[-3, 4, -1] + \alpha_3[6, -8, 2] = [0, 0, 0]$$

By comparing the corresponding components, we have

$$\begin{aligned} 3\alpha_1 - 3\alpha_2 + 6\alpha_3 &= 0 \\ -4\alpha_1 + 4\alpha_2 - 8\alpha_3 &= 0 \\ \alpha_1 - \alpha_2 + 2\alpha_3 &= 0 \end{aligned} \qquad \text{(i)}$$

The matrix form of the system (i) is

$$\mathbf{Ax} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 3 & -3 & 6 \\ -4 & 4 & -8 \\ 1 & -1 & 2 \end{bmatrix}; \qquad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

The augmented matrix is

$$\begin{bmatrix} 3 & -3 & 6 & : & 0 \\ -4 & 4 & -8 & : & 0 \\ 1 & -1 & 2 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 & : & 0 \\ 0 & 0 & 0 & : & 0 \\ 1 & -1 & 2 & : & 0 \end{bmatrix}$$

$$\alpha_1 - \alpha_2 + 2\alpha_3 = 0$$

Therefore the values of α_1, α_2 and α_3 depend on each other, so the given three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ do lie on the same line.

(ii) For $\mathbf{v}_1 = [3, 2, -1]$, $\mathbf{v}_2 = [-6, -4, 2]$, $\mathbf{v}_3 = [0, 6, -3]$
the vector equation is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2 \text{ and } \alpha_3$$

$$\alpha_1[3, 2, -1] + \alpha_2[-6, -4, 2] + \alpha_3[0, 6, -3] = [0, 0, 0]$$

By comparing the corresponding components, we have

$$\begin{aligned} 3\alpha_1 - 6\alpha_2 \qquad &= 0 \\ 2\alpha_1 - 4\alpha_2 + 6\alpha_3 &= 0 \\ -\alpha_1 + 2\alpha_2 - 3\alpha_3 &= 0 \end{aligned} \tag{i}$$

The matrix forms of the system (i) is

$$\mathbf{Ax} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 3 & -6 & 0 \\ 2 & -4 & 6 \\ -1 & 2 & -3 \end{bmatrix}; \qquad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

The augmented matrix is

$$\begin{bmatrix} 3 & -6 & 0 & : & 0 \\ 2 & -4 & 6 & : & 0 \\ -1 & 2 & -3 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & -9 & : & 0 \\ 0 & 0 & 0 & : & 0 \\ -1 & 2 & -3 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 1 & : & 0 \\ 0 & 0 & 0 & : & 0 \\ -1 & 2 & 0 & : & 0 \end{bmatrix}$$

$$\alpha_1 = 2\alpha_2, \ \alpha_3 = 0$$

This shows that the values of α_1 and α_2 depend on each other but not on α_3, so all three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ do not lie on the same line but $\mathbf{v}_1$ and $\mathbf{v}_2$ do lie on the same line passing through the origin.

(iii) For $\mathbf{v}_1 = [1, 4, 1]$, $\mathbf{v}_2 = [3, 1, -1]$, $\mathbf{v}_3 = [4, 5, 0]$
the vector equation is

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2 \text{ and } \alpha_3$$

$$\alpha_1[1, 4, 1] + \alpha_2[3, 1, -1] + \alpha_3[4, 5, 0] = [0, 0, 0]$$

By comparing the corresponding components, we have

$$\begin{aligned} \alpha_1 + 3\alpha_2 + 4\alpha_3 &= 0 \\ 4\alpha_1 + \alpha_2 + 5\alpha_3 &= 0 \\ \alpha_1 - \alpha_2 \qquad &= 0 \end{aligned}$$

The matrix forms of the system (i) is

$$\mathbf{Ax} = \mathbf{0}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 4 & 1 & 5 \\ 1 & -1 & 0 \end{bmatrix}; \qquad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

The augmented matrix is

$$\begin{bmatrix} 1 & 3 & 4 & : & 0 \\ 4 & 1 & 5 & : & 0 \\ 1 & -1 & 0 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 4 & 4 & : & 0 \\ 0 & 5 & 5 & : & 0 \\ 1 & -1 & 0 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 1 & -1 & 0 & : & 0 \end{bmatrix}$$

$$\alpha_1 - \alpha_2 = 0, \ \alpha_2 + \alpha_3 = 0$$

$$\alpha_1 = \alpha_2 = -\alpha_3$$

This is the equation of a plane in R^3, that is, $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are not multiple of each other.

Linear Independence of Functions

Consider the set $F(-\infty, \infty)$ of real-valued functions. In Example 2.37, we used some identities of trigonometric functions to prove the linear independence of functions. Although in general, we do not have any standard method to establish linear independence or linear dependence between functions in $F(-\infty, \infty)$, the theorem stated below can be used to show the linear independence of functions of $F(-\infty, \infty)$.

Let us start with definition.

Definition: Wronskian

Let $f_1(x), f_2(x), \ldots, f_n(x)$ be $(n - 1)$ times differentiable functions of $F(-\infty, \infty)$. Then the determinant,

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1^1(x) & f_2^1(x) & \cdots & f_n^1(x) \\ . & . & \cdots & . \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \cdots & f_n^{n-1}(x) \end{vmatrix}$$

is called Wronskian of $f_1(x), f_2(x), \ldots, f_n(x)$.

Theorem 2.12 [Linear Independence of Functions]

If the functions $f_1, f_2, \ldots, f_n$ have continuous derivatives on the interval $(-\infty, \infty)$ and if the Wronskian $W(x)$ of these functions is not identically zero on $(-\infty, \infty)$, then these functions form a linearly independent set of vectors in $C^{n-1}(-\infty, \infty)$.

EXAMPLE 2.46 Which of the following sets of vectors in $F(-\infty, \infty)$ are linearly independent?

(i) $x, \cos x$
(ii) $\cos 2x, \sin^2 x, \cos^2 x$
(iii) $0, x^2 - 6x, 5$
(iv) e^x, xe^x, x^2e^x

Solution:

(i) Say

$$f_1(x) = x, \qquad f_2(x) = \cos x$$
$$f_1'(x) = 1, \qquad f_2'(x) = -\sin x$$

The Wronskian $W(x)$ of these functions is

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_1'(x) \end{vmatrix}$$
$$= \begin{vmatrix} x & \cos x \\ 1 & -\sin x \end{vmatrix}$$
$$= -x \sin x - \cos x$$

If $x = 0$, then $W(x) = -1$. So this function $W(x)$ does not have a value zero for all x in the interval $(-\infty, \infty)$. Hence f_1 and f_2 form a linearly independent set.

(ii) Say

$$f_1(x) = \cos 2x, \qquad f_2(x) = \sin^2 x, \qquad f_3(x) = \cos^2 x$$
$$f_1'(x) = -2\sin 2x, \qquad f_2'(x) = \sin 2x, \qquad f_3'(x) = -\sin 2x$$
$$f_1''(x) = -4\cos 2x, \qquad f_2''(x) = 2\cos 2x, \qquad f_3''(x) = -2\cos 2x$$

The Wronskian $W(x)$ of these functions is

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) \end{vmatrix}$$
$$= \begin{vmatrix} \cos 2x & \sin^2 x & \cos^2 x \\ -2\sin 2x & \sin 2x & -\sin 2x \\ -4\cos 2x & 2\cos 2x & -2\cos 2x \end{vmatrix}$$
$$= \cos 2x\,(-2\cos 2x \sin 2x + 2\sin 2x \cos 2x)$$
$$-\sin^2 x\,(4\sin 2x\cos 2x - 4\sin 2x \cos 2x)$$
$$+\cos^2 x\,(-4\sin 2x \cos 2x + 4\sin 2x\cos 2x)$$
$$= 0$$

Therefore the Wronskian is identically zero. So, f_1, f_2 and f_3 form a linearly dependent set.

(iii) Say

$$\begin{aligned} &f_1(x) = 0, && f_2(x) = x^2 - 6x, && f_3(x) = 5 \\ &f_1'(x) = 0, && f_2'(x) = 2x - 6, && f_3'(x) = 0 \\ &f_1''(x) = 0, && f_2''(x) = 2, && f_3''(x) = 0 \end{aligned}$$

The Wronskian $W(x)$ of these functions is

$$\begin{aligned} W(x) &= \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) \end{vmatrix} \\ &= \begin{vmatrix} 0 & x^2-6x & 5 \\ 0 & 2x-6 & 0 \\ 0 & 2 & 0 \end{vmatrix} \\ &= 0 \end{aligned}$$

Therefore the Wronskian is identically zero. So, f_1, f_2 and f_3 form a linearly dependent set.

(iv) Say

$$\begin{aligned} &f_1(x) = e^x, && f_2(x) = xe^x, && f_3(x) = x^2e^x \\ &f_1'(x) = e^x, && f_2'(x) = (x+1)e^x, && f_3'(x) = (x^2+2x)e^x \\ &f_1''(x) = e^x, && f_2''(x) = (x+2)e^x, && f_3''(x) = (x^2+4x+2)e^x \end{aligned}$$

The Wronskian $W(x)$ of these functions is

$$\begin{aligned} W(x) &= \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) \end{vmatrix} \\ &= \begin{vmatrix} e^x & xe^x & x^2e^x \\ e^x & (x+1)e^x & (x^2+2x)e^x \\ e^x & (x+2)e^x & (x^2+4x+2)e^x \end{vmatrix} \\ &= e^{3x}\begin{vmatrix} 1 & x & x^2 \\ 1 & x+1 & x^2+2x \\ 1 & x+2 & x^2+4x+2 \end{vmatrix} \\ &= e^{3x}\begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 2 & 4x+2 \end{vmatrix} \\ &= e^{3x}(4x + 2 - 4x) = 2e^{3x} \neq 0 \end{aligned}$$

This function $W(x)$ does not have the value zero for any x in the interval $(-\infty, \infty)$. Hence, f_1, f_2 and f_3 form a linearly independent set.

EXERCISE SET 3

1. Examine the linear dependency and linear independency of the following sets of vectors.

(i) [1, 0, 2], [1, –1, –1], [3, 2, 4]

(ii) [1, 1, 0], [0, 1, 1], [1, 0, 1]

(iii) [2, 0, 0], [0, 2, 0], [0, 0, 2]

(iv) [2, 1, –1], [1, –7, –8], [1, 3, 2]

(v) [2, 1, 4], [1, –1, 2], [3, 1, –2]

(vi) [1, 2, 1, 0], [1, 3, 1, 2], [4, 2, 1, 0], [6, 1, 0, 1]

(vii) [1, 1, 1], [5, –1, 1], [0, 0, 0]

2. Which of the following sets of polynomials are linearly independent and which linearly dependent?

(i) $\{1 + 3x, 5x, 1 + 5x\}$ in P_1.

(ii) $\{1 + 2x + 5x^2, 3 + 6x + 5x^2\}$ in P_2.

(iii) $\{-3 + 4x^2, 5, 5 - x + 2x^2, 1 + x + 5x^2\}$ in P_2.

(iv) $\{1 + 3x + x^2 + 2x^3, x, 2x - 2x^2, -1 + 2x + x^2 + x^3\}$ in P_3.

3. Which of the following functions are linearly independent and which linearly dependent?

(i) $x, \sin x$ (ii) $\log x, -\log x^4, \log x^5$ (iii) $e^x \cos x, e^x \sin x, e^x$

(iv) $(x - a)^2, (x + a), x$ (v) $0, 1, x, x^2$.

2.4 BASIS AND DIMENSION

In this section, we will complete the task of describing uniquely, every element of a vector space V in terms of the elements of a suitable set S of V. We will begin this section with the following definition.

Definition: Basis Vectors

Suppose $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors from the space V. Then S is called a basis for V if both the following conditions hold.

(i) S spans the vector space V, that is, $V = \text{span}(S)$.

(ii) S is a linearly independent set of vectors.

Now we first consider the standard basis for same vector spaces.

EXAMPLE 2.47 Show that the following sets of vectors are bases for the indicated vector spaces.

(i) $\mathbf{i} = [1, 0], \mathbf{j} = [0, 1]$ for R^2

(ii) $\mathbf{i} = [1, 0, 0], \mathbf{j} = [0, 1, 0], \mathbf{k} = [0, 0, 1]$ for R^3

(iii) $\mathbf{e}_1 = [1, 0, 0, \ldots, 0], \mathbf{e}_2 = [0, 1, 0, \ldots, 0], \ldots, \mathbf{e}_n = [0, 0, 0, \ldots, 1]$

Solution:

(i) Let $S = \{\mathbf{i}, \mathbf{j}\}$

To prove that S is a basis for R^2, we have to show that S is a linearly independent set and spans R^2.

Consider the vector equation,

$$\alpha_1\mathbf{i} + \alpha_2\mathbf{j} = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2$$

$$\alpha_1[1, 0] + \alpha_2[0, 1] = [0, 0]$$
$$[\alpha_1, \alpha_2] = [0, 0]$$
$$\alpha_1 = 0, \ \ \alpha_2 = 0$$

Therefore the set S is a linearly independent set.
Generalizing:
Let $[a, b]$ be any vector of R^2.
Suppose $[a, b] = c_1\,\mathbf{i} + c_2\,\mathbf{j}$ for some scalars c_1, c_2.

$$[a, b] = c_1[1, 0] + c_2[0, 1]$$
$$[a, b] = [c_1, c_2]$$
$$c_1 = a, \ \ c_2 = b$$

Thus $[a, b] = a\mathbf{i} + b\mathbf{j}$. Therefor S spans R^2.
Thus S is a basis for R^2 and it is known as the *standard basis* of R^2.

(ii) Let $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$
To prove that S is a basis for R^3, we have to show that S is linearly independent and spans R^3.
Consider the vector equation,

$$\alpha_1\mathbf{i} + \alpha_2\mathbf{j} + \alpha_3\mathbf{k} = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3$$
$$\alpha_1[1, 0, 0] + \alpha_2[0, 1, 0] + \alpha_3[0, 0, 1 \ = [0, 0, 0]$$
$$[\alpha_1, \alpha_2, \alpha_3] \ = [0, 0, 0]$$
$$\alpha_1 = \alpha_2 = \alpha_3 = 0$$

Therefore the set S is linearly independent.
Generalizing:
Let $[a, b, c]$ be any vector of R^3.
Suppose $\quad [a, b, c] = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$ for some scalars c_1, c_2, c_3.

$$[a, b, c] = c_1[1, 0, 0] + c_2[0, 1, 0] + c_3[0, 0, 1]$$
$$[a, b, c] = [c_1, c_2, c_3]$$
$$c_1 = a, \ \ c_2 = b, \ \ c_3 = c$$

Thus $[a, b, c] = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. Therefore S spans R^3.
Thus S is a basis for R^3 and it is known as the *standard basis* of R^3.

(iii) Let $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$
Consider the vector equation

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, ..., \alpha_n$$
$$\alpha_1[1, 0, 0, \ldots, 0] + \alpha_2[0, 1, 0, \ldots, 0] + \cdots + \alpha_n[0, 0, \ldots, 1] = [0, 0, \ldots, 0]$$
$$\therefore \qquad [\alpha_1, \alpha_2, \ldots, \alpha_n] = [0, 0, \ldots, 0]$$

Hence $\qquad \alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_n = 0$

Therefore the set S is linearly independent.
Generalizing:
Let $[a_1, a_2, \ldots, a_n]$ be any vector of R^n. Suppose

$$[a_1, a_2, \ldots, a_n] = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n \qquad \text{for some scalars } c_1, c_2, \ldots, c_n.$$

$$[a_1, a_2, \ldots, a_n] = c_1[1, 0, 0, \ldots, 0] + c_2[0, 1, 0, \ldots, 0] + \cdots + c_n\,[0, 0, 0, \ldots, 1]$$

$\therefore \quad [a_1, a_2, \ldots, a_n] = [c_1, c_2, \ldots, c_n]$

Hence $\quad a_1 = c_1,\ a_2 = c_2,\ \ldots,\ a_n = c_n$

Thus $\quad [a_1, a_2, \ldots, a_n] = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n$

Therefore S spans R^n.

Thus S is a basis for R^n and it is known as the *standard basis* of R^n.

EXAMPLE 2.48 Show that the following sets of vector are bases for the indicated vector spaces.

(i) $\mathbf{M}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{M}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{M}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{M}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ for M_{22}

(ii) $S = \{1, x, x^2, x^3, \ldots, x^n\}$ for P_n.

Solution:

(i) Let $M_{22} = \{\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4\}$

Consider the vector equation,

$$\alpha_1\mathbf{M}_1 + \alpha_2\mathbf{M}_2 + \alpha_3\mathbf{M}_3 + \alpha_4\mathbf{M}_4 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3, \alpha_4.$$

$$\alpha_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \alpha_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\alpha_1 = 0,\ \alpha_2 = 0,\ \alpha_3 = 0,\ \alpha_4 = 0.$$

Therefore S is a linearly independent set.

Generalizing:

Let $\mathbf{M} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ be any vector of M_{22}.

Suppose $\mathbf{M} = c_1\mathbf{M}_1 + c_2\mathbf{M}_2 + c_3\mathbf{M}_3 + c_4\mathbf{M}_4$ for some scalars c_1, c_2, c_3, c_4.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = c_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\therefore \quad \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$$

Hence $\quad c_1 = a_1,\ c_2 = a_2,\ c_3 = a_3,\ c_4 = a_4.$

Thus $\mathbf{M} = a_1\mathbf{M}_1 + a_2\mathbf{M}_2 + a_3\mathbf{M}_3 + a_4\mathbf{M}_4$. Therefore S spans M_{22}.

Thus S is a basis for M_{22} and it is known as the *standard basis* of M_{22}.

(ii) Let $S = \{1, \mathbf{x}, \mathbf{x}^2, \mathbf{x}^3, \ldots, \mathbf{x}^n\}$ for P_n.

Consider the vector equation,

$$\alpha_0 + \alpha_1\mathbf{x} + \alpha_2\mathbf{x}^2 + \alpha_3\mathbf{x}^3 + \cdots + \alpha_n\mathbf{x}^n = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \ldots, \alpha_n$$

$\therefore \quad \alpha_0 + \alpha_1\mathbf{x} + \alpha_2\mathbf{x}^2 + \alpha_3\mathbf{x}^3 + \cdots + \alpha_n\mathbf{x}^n = 0 + 0\mathbf{x} + 0\mathbf{x}^2 + \cdots + 0\mathbf{x}^n$

Hence $\quad \alpha_0 = 0, \ \alpha_1 = 0, \ \ldots, \ \alpha_n = 0$

Therefore S is a linearly independent set.

Since each polynomial matrix $p(\mathbf{x})$ in P_n can be written as $p(\mathbf{x}) = a_0\mathbf{I} + a_1\mathbf{x} + \cdots + a_n\mathbf{x}^n$ which is a linear combination of $\mathbf{I}, \mathbf{x}, \ldots, \mathbf{x}^n$. Therefore S spans P_n.

Thus S is basis for P_n and it is known as the *standard basis* for P_n.

The most important and unique property of the basis of vector spaces is described in the following theorem.

Theorem 2.13 [Basis]

Suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for vector space V. Then every vector $\mathbf{u} \in V$ can be expressed uniquely in the form.

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \qquad \text{where } c_1, c_2, \ldots, c_n \text{ are scalars.}$$

EXAMPLE 2.49 Which of the following sets of vectors are bases for R^2?

(i) $S_1 = \{[1, 1], [-1, 1]\}$ (ii) $S_2 = \{[6, -2], [-3, 1]\}$ (iii) $S_3 = \{[1, 3], [0, 0]\}$

Solution:

(i) Let $\mathbf{v}_1 = [1, 1]$, $\mathbf{v}_2 = [-1, 1]$

Consider the vector equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0} \qquad \text{where } \alpha_1, \alpha_2 \text{ are scalars.}$$

$$\alpha_1[1, 1] + \alpha_2[-1, 1] = [0, 0]$$

$$[\alpha_1 - \alpha_2, \alpha_1 + \alpha_2] = [0, 0]$$

$$\alpha_1 - \alpha_2 = 0, \ \alpha_1 + \alpha_2 = 0$$

$$\alpha_1 = 0, \ \alpha_2 = 0$$

Therefore S_1 is a linearly independent set.

Generalizing:

Let $[a, b]$ be any point of R^2.

Suppose

$$[a, b] = c_1[1, 1] + c_2[-1, 1]$$

$\therefore \quad [a, b] = [c_1 - c_2, c_1 + c_2]$

or $\quad a = c_1 - c_2, \ b = c_1 + c_2$

Hence $\quad c_1 = \dfrac{a+b}{2}, \ c_2 = \dfrac{a-b}{2}.$

Therefore S_1 spans R^2.

Hence S_1 is a basis for R^2.

(ii) Let $\mathbf{v}_1 = [6, -2]$, $\mathbf{v}_2 = [-3, 1]$

Consider the vector equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2.$$

$$\alpha_1[6, -2] + \alpha_2[-3, 1] = [0, 0]$$

$\therefore$ $$[6\alpha_1 - 3\alpha_2, -2\alpha_1 + \alpha_2] = [0, 0]$$

Hence $$6\alpha_1 - 3\alpha_2 = 0, -2\alpha_1 + \alpha_2 = 0$$

or $$2\alpha_1 = \alpha_2.$$

Thus $$\alpha_1\mathbf{v}_1 + 2\alpha_1\mathbf{v}_2 = \mathbf{0}$$

$\therefore$ $$\mathbf{v}_1 = -2\mathbf{v}_2$$

Therefore S_2 is a linearly dependent set.
Thus S_2 cannot be a basis for R^2

(iii) Since the set S_3 contains zero vectors [0, 0], so S_3 is a linearly dependent set. Hence S_3 cannot be a basis for R^2.

EXAMPLE 2.50 Which of the following sets of vectors are bases for R^3?

(i) $S_1 = \{[1, 0, 0], [2, 2, 0], [3, 3, 3]\}$
(ii) $S_2 = \{[1, -2, 3], [5, 6, -1], [3, 2, 1]\}$

Solution:

(i) Let $\mathbf{v}_1 = [1, 0, 0]$, $\mathbf{v}_2 = [2, 2, 0]$, $\mathbf{v}_3 = [3, 3, 3]$
Consider the vector equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3.$$

$$\alpha_1[1, 0, 0] + \alpha_2[2, 2, 0] + \alpha_3[3, 3, 3] = [0, 0, 0]$$

$$\alpha_1 + 2\alpha_2 + 3\alpha_3 = 0$$

$$2\alpha_2 + 3\alpha_3 = 0$$

$$3\alpha_3 = 0$$

The matrix of coefficients of the above system is

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{vmatrix} = 6 \neq 0$$

It has a nonzero determinant. So S_1 is a linearly independent set.
Let $[a, b, c]$ be any vector in R^3.

$$[a, b, c] = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \qquad \text{for some scalars } c_1, c_2, c_3.$$

$$[a, b, c] = c_1[1, 0, 0] + c_2[2, 2, 0] + c_3[3, 3, 3]$$

$$a = c_1 + 2c_2 + 3c_3$$

$$b = 2c_2 + 3c_3$$

$$c = 3c_3$$

Thus $$c_3 = \frac{c}{3},\ c_2 = \frac{b-c}{2},\ c_1 = a-b$$

Therefore $$(a, b, c) = (a-b)\,\mathbf{v}_1 + \frac{b-c}{2}\,\mathbf{v}_2 + \frac{c}{3}\,\mathbf{v}_3.$$

Thus S_1 spans R^3. Hence S_1 is a basis for R^3.

(ii) $\mathbf{v}_1 = [1, -2, 3]$, $\mathbf{v}_2 = [5, 6, -1]$, $\mathbf{v}_3 = [3, 2, 1]$

Consider the vector equation,

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3.$$

$$\alpha_1[1, -2, 3] + \alpha_2[5, 6, -1] + \alpha_3[3, 2, 1] = [0, 0, 0]$$

$$\begin{aligned} \alpha_1 + 5\alpha_2 + 3\alpha_3 &= 0 \\ -2\alpha_1 + 6\alpha_2 + 2\alpha_3 &= 0 \\ 3\alpha_1 - \alpha_2 + \alpha_3 &= 0 \end{aligned}$$

The matrix of coefficient of the above system is

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 1 & 5 & 3 \\ -2 & 6 & 2 \\ 3 & -1 & 1 \end{vmatrix} = 0$$

It has a zero determinant. So S_2 is a linearly dependent set.
Therefore S_2 cannot be a basis for R^3.

EXAMPLE 2.51 Which of the following sets of vectors are bases for P_2?

(i) $S_1 = \{1 + 2x + x^2, 2 + 9x, 3 + 3x + 4x^2\}$

(ii) $S_2 = \{2 - 3x + x^2, 4 + x + x^2, 7x + x^2\}$

Solution:

(i) Let $\mathbf{p}_1 = 1 + 2x + x^2$, $\mathbf{p}_2 = 2 + 9x$, $\mathbf{p}_3 = 3 + 3x + 4x^2$

Consider the vector equation,

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3.$$

$$\alpha_1(1 + 2x + x^2) + \alpha_2(2 + 9x) + \alpha_3(3 + 3x + 4x^2) = \mathbf{0}$$

$$\begin{aligned} \alpha_1 + 2\alpha_2 + 3\alpha_3 &= 0 \\ 2\alpha_1 + 9\alpha_2 + 3\alpha_3 &= 0 \\ \alpha_1 \qquad\quad + 4\alpha_3 &= 0 \end{aligned}$$

The matrix of the coefficients of the above system is

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix} \qquad \Rightarrow \qquad \det \mathbf{A} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{vmatrix} = -1 \neq 0$$

Therefore S_1 is a linearly independent set of vectors of P_2.

Generalizing:

Let $\mathbf{p} = b_0 + b_1x + a_2x^2$ be any vector of P_2.

Suppose $\mathbf{p} = c_1\mathbf{p}_1 + c_2\mathbf{p}_2 + c_3\mathbf{p}_3$ for some scalars c_1, c_2, c_3.

$$b_0 + b_1x + b_2x^2 = c_1(1 + 2x + x^2) + c_2(2 + 9x) + c_3(3 + 3x + 4x^2)$$

$$\begin{aligned} c_1 + 2c_2 + 3c_3 &= b_0 \\ 2c_1 + 9c_2 + 3c_3 &= b_1 \\ c_1 \qquad + 4c_3 &= b_2 \end{aligned}$$

The augmented matrix of the system is

$$\begin{bmatrix} 1 & 2 & 3 & : & b_0 \\ 2 & 9 & 3 & : & b_1 \\ 1 & 0 & 4 & : & b_2 \end{bmatrix}$$

By using the Gauss-elimination method, we have

$$\begin{bmatrix} 1 & 0 & 0 & : & -36b_0 + 8b_1 + 21b_2 \\ 0 & 1 & 0 & : & 5b_0 - b_1 - 3b_2 \\ 0 & 0 & 1 & : & 9b_0 - 2b_1 - 5b_2 \end{bmatrix}$$

$$\begin{aligned} c_1 &= -36b_0 + 8b_1 + 2b_2 \\ c_2 &= 5b_0 - b_1 - 3b_2 \\ c_3 &= 9b_0 - 2b_1 - 5b_2 \end{aligned}$$

Therefore S_1 spans P_2.

Thus S_1 is a basis for P_2.

(ii) Let $\mathbf{p}_1 = 2 - 3x + x^2$, $\mathbf{p}_2 = 4 + x + x^2$, $\mathbf{p}_3 = -7x + x^2$

Consider the vector equation,

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 = \mathbf{0}$$

$$\alpha_1(2 - 3x + x^2) + \alpha_2(4 + x + x^2) + \alpha_3(-7x + x^2) = 0$$

$$\begin{aligned} 2\alpha_1 + 4\alpha_2 \qquad &= 0 \\ -3\alpha_1 + \alpha_2 - 7\alpha_3 &= 0 \\ \alpha_1 + \alpha_2 + \alpha_3 &= 0 \end{aligned}$$

The matrix of the coefficients of the above system is

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 2 & 4 & 0 \\ -3 & 1 & -7 \\ 1 & 1 & 1 \end{vmatrix} = 16 - 16 = 0$$

It has a zero determinant. So S_2 is a linearly independent set. Therefore S_2 cannot be a basis for P_2.

EXAMPLE 2.52 If $\mathbf{S} = \left\{ \begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \right\}$, then show that $\mathbf{S}$ is a basis for M_{22}.

Solution: Let $\mathbf{M}_1 = \begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}$, $\mathbf{M}_2 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, $\mathbf{M}_3 = \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}$, $\mathbf{M}_4 = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

The vector equation for $\mathbf{S}$ is

$$\alpha_1\mathbf{M}_1 + \alpha_2\mathbf{M}_2 + \alpha_3\mathbf{M}_3 + \alpha_4\mathbf{M}_4 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3, \alpha_4.$$

$$\alpha_1 \begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix} + \alpha_4 \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{aligned} 3\alpha_1 \qquad\qquad\qquad + \alpha_4 &= 0 \\ 6\alpha_1 - \alpha_2 - 8\alpha_3 \qquad\quad &= 0 \\ 3\alpha_1 - \alpha_2 - 12\alpha_3 - \alpha_4 &= 0 \\ -6\alpha_1 \qquad - 4\alpha_3 + 2\alpha_4 &= 0 \end{aligned}$$

The matrix of the coefficients of the above equation is

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 6 & -1 & -8 & 0 \\ 3 & -1 & -12 & -1 \\ -6 & 0 & -4 & 2 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 3 & 0 & 0 & 1 \\ 6 & -1 & -8 & 0 \\ 3 & -1 & -12 & -1 \\ -6 & 0 & -4 & 2 \end{vmatrix} = 3\begin{vmatrix} -1 & -8 & 0 \\ -1 & -12 & -1 \\ 0 & -4 & 2 \end{vmatrix} - 1\begin{vmatrix} 6 & -1 & -8 \\ 3 & -1 & -12 \\ -6 & 0 & -4 \end{vmatrix} = 48 \neq 0$$

Therefore S is a linearly independent set.
Generalizing:

Let $\mathbf{M} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ be any vector of M_{22}.

Suppose $\mathbf{M} = c_1\mathbf{M}_1 + c_2\mathbf{M}_2 + c_3\mathbf{M}_3 + c_4\mathbf{M}_4$ for some scalars c_1, c_2, c_3, c_4.

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = c_1\begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix} + c_2\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix} + c_4\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{aligned} 3c_1 \qquad\qquad\qquad + \ c_4 &= a_1 \\ 6c_1 - c_2 - \ 8c_3 \qquad\quad &= a_2 \\ 3c_1 - c_2 - 12c_3 - \ c_4 &= a_3 \\ -6c_1 \qquad - \ 4c_3 + 2c_4 &= a_4 \end{aligned}$$

The augmented matrix of the system is

$$\begin{bmatrix} 3 & 0 & 0 & 1 & : & a_1 \\ 6 & -1 & -8 & 0 & : & a_2 \\ 3 & -1 & -12 & -1 & : & a_3 \\ -6 & 0 & -4 & 2 & : & a_4 \end{bmatrix}$$

By using the Gauss-elimination method,

$$\begin{bmatrix} 12 & 0 & 0 & 0 & : & 3a_1 - a_2 + a_3 - a_4 \\ 0 & -2 & 0 & 0 & : & 7a_2 - 7a_1 - 5a_3 + a_4 \\ 0 & 0 & -4 & 0 & : & a_1 - a_2 + a_3 \\ 0 & 0 & 0 & 4 & : & a_1 + a_2 - a_3 + a_4 \end{bmatrix}$$

$$c_1 = \frac{3a_1 - a_2 + a_3 - a_4}{12}$$

$$c_2 = \frac{-7a_2 + 7a_1 + 5a_3 - a_4}{2}$$

$$c_3 = \frac{a_2 - a_1 - a_3}{4}$$

$$c_4 = \frac{a_1 + a_2 - a_3 + a_4}{4}$$

Thus S spans M_{22}. Therefore S is a basis for M_{22}.

Co-ordinates Relatives to a Basis

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, and

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \ldots, c_n$ are called the co-ordinates of $\mathbf{v}$ relative to the basis S. It is denoted by $(\mathbf{v})_S = [c_1, c_2, \ldots, c_n]$.

EXAMPLE 2.53 Find the co-ordinate vector of $\mathbf{v} = [1, 0]$ relative to the basis $S = \{[1, -1], [1, 1]\}$.

Solution: Let $\mathbf{v}_1 = [1, -1]$, $\mathbf{v}_2 = [1, 1]$
We want to find scalars c_1, c_2 such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$$

$$[1, 0] = c_1\,[1, -1] + c_2\,[1, 1]$$
$$[1, 0] = [c_1 + c_2, -c_1 + c_2]$$
$$1 = c_1 + c_2, \;\; 0 = -c_1 + c_2$$
$$c_1 = \frac{1}{2}, \qquad c_2 = \frac{1}{2}$$

Therefore the co-ordinate vector relative to S is $[\mathbf{v}]_S = \left[\frac{1}{2}, \frac{1}{2}\right]$.

EXAMPLE 2.54 Find the co-ordinate vector of $\mathbf{v} = [2, -1, 1]$ relative to the basis $S = \{[1, 1, 0], [1, 0, 1], [0, 1, 1]\}$.

Solution: Let $\mathbf{v}_1 = [1, 1, 0]$, $\mathbf{v}_2 = [1, 0, 1]$, $\mathbf{v}_3 = [0, 1, 1]$.

We want to find scalars c_1, c_2, c_3 such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$
$$[2, -1, 1] = c_1\,[1, 1, 0] + c_2[1, 0, 1] + c_3[0, 1, 1]$$
$$c_1 + c_2 = 2$$
$$c_1 + c_3 = -1$$
$$c_2 + c_3 = 1$$

By solving this system, we get $c_1 = 0$, $c_2 = 2$, $c_3 = -1$. Therefore the co-ordinate vector of $\mathbf{v}$ relative to S is

$$[\mathbf{v}]_S = [c_1, c_2, c_3] = [0, 2, -1]$$

EXAMPLE 2.55 Find the co-ordinate vector of $\mathbf{p} = [2 - x + 3x^2]$ relative to the basis $S = \{[1], 2 + 2x], [3 + 3x + 3x^2]\}$.

Solution: Let $\mathbf{p}_1 = 1$, $\mathbf{p}_2 = 2 + 2x$, $\mathbf{p}_3 = [3 + 3x + 3x^2]$

We have to find the scalars c_1, c_2, c_3 such that

$$\mathbf{p} = c_1\mathbf{p}_1 + c_2\mathbf{p}_2 + c_3\mathbf{p}_3$$
$$[2 - x + 3x^2] = c_1[1] + c_2\,[2 + 2x] + c_3\,[3 + 3x + 3x^2]$$
$$c_1 + 2c_2 + 3c_3 = 2$$
$$2c_2 + 3c_3 = -1$$
$$3c_3 = 3$$

By solving this system, we get $c_1 = 3$, $c_2 = -2$, $c_3 = 1$.

Therefore the co-ordinate vector of $\mathbf{p}$ relative to S is

$$[\mathbf{p}]_S = [c_1, c_2, c_3] = [3, -2, 1]$$

EXAMPLE 2.56 Find the co-ordinates vector of $\mathbf{M} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}$ relative to the basis

$$S = \left\{\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}.$$

Solution: Let $\mathbf{M}_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{M}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{M}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\mathbf{M}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

We want to find scalars c_1, c_2, c_3, c_4 such that

$$\mathbf{M} = c_1\mathbf{M}_1 + c_2\mathbf{M}_2 + c_3\mathbf{M}_3 + c_4\mathbf{M}_4$$

$$\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} = c_1\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$-c_1 + c_2 = 2,\ c_1 + c_2 = 0,\ c_3 = -1,\ c_4 = 3$$

Thus $c_1 = -1$, $c_2 = 1$, $c_3 = -1$, $c_4 = 3$. Therefore the co-ordinates vector of $\mathbf{M}$ relative to S is

$$[\mathbf{M}]_S = [-1, 1, -1, 3].$$

Dimensions

We have shown earlier that a vector space can have many bases. For example, a collection of three vectors not on the same plan is a basis for R^3. Now in the following discussion, we will attempt to find out some more properties of bases. Let us start with some definitions of dimension of vector space.

Definitions: Dimension of Vector Space

1. A vector space V over R is said to be *finite-dimensional* if it has a basis containing only finitely many elements, otherwise it is called an *infinite-dimensional* vector space. For example, as seen in Examples 2.47 and 2.48 the vector spaces R^n, M_{22} and P_n are finite-dimensional.
2. The vector spaces $F(-\infty, \infty)$ of real-valued functions defined in the interval $(-\infty, \infty)$ are the examples of infinite-dimensional vector spaces.

EXAMPLE 2.57 Let $S = \{[1, 0, 1], [0, 1, 1], [1, 1, 0]\}$ be a basis for R^3. Are the following sets of vectors bases for R^3?

(i) $\mathbf{u}_1 = [1, 0, 1]$, $\mathbf{u}_2 = [0, 1, 1]$
(ii) $\mathbf{u}_1 = [1, 0, 1]$, $\mathbf{u}_2 = [0, 1, 1]$, $\mathbf{u}_3 = [1, 1, 0]$, $\mathbf{u}_4 = [1, 1, 1]$

Solution:

(i) Let $S_1 = \{\mathbf{u}_1, \mathbf{u}_2\}$

Consider the vector equation

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2$$

$$\alpha_1[1, 0, 1] + \alpha_2[0, 1, 1] = [0, 0, 0]$$

$$\alpha_1 = 0,\ \alpha_2 = 0$$

Therefore S is a linearly independent set of R^3.

Generalizing:

Let $\mathbf{u} = [a, b, c]$ be any vector of R^3.

$$\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$$

$$[a, b, c] = \alpha_1[1, 0, 1] + \alpha_2[0, 1, 1]$$

$$[a, b, c] = [\alpha_1, \alpha_2, \alpha_1 + \alpha_2]$$

$$a = \alpha_1,\ b = \alpha_2,\ c = \alpha_1 + \alpha_2.$$

$$\therefore \qquad \alpha_1 = a,\ \alpha_2 = b,\ a + b = c$$

Thus the vector $\mathbf{u} = [a, b, c]$ can be expressed as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$, only when $a + b = c$, that is, all the vectors of R^3 cannot be expressed as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$. So the given set S is not a basis for R^3.

(ii) Let $S_2 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$
Consider the vector equation

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3, \alpha_4.$$

$$\alpha_1[1, 0, 1] + \alpha_2[0, 1, 1] + \alpha_3[1, 1, 0] + \alpha_4[1, 1, 1] = [0, 0, 0]$$

$$\alpha_1 + \alpha_3 + \alpha_4 = 0$$

$$\alpha_2 + \alpha_3 + \alpha_4 = 0$$

$$\alpha_1 + \alpha_2 + \alpha_4 = 0$$

Solving this system, we obtain $\alpha_1 = \frac{1}{2}, \alpha_2 = \frac{1}{2}, \alpha_3 = \frac{1}{2}, \alpha_4 = -1$. That means the system has a non-trivial solution. So S_2 is a linearly dependent set. Therefore S_2 is not a basis for R^3.

By generalizing the result of Example 2.57, we have the following theorems.

Theorem 2.14 [Finite Dimensional Vector Space]

Let V be a finite-dimensional vector space and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis for V. Then:

(i) If the set has more than n vectors, it is linearly dependent.
(ii) If the set has fewer than n vectors, it does not span V.

Theorem 2.15 [Basis for a Vector Space]

Any two bases for a vector space must contain the same number of elements. This theorem suggests the following definition.

Definition: Dimension of Vector Space

The number of vectors in a basis for a finite-dimensional vector space V is called the dimension of V. It is denoted by

$$\dim V.$$

For example:

(i) The basis $\{[1, 0, 0, \ldots, 0], [0, 1, 0, \ldots, 0], \ldots, [0, 0, \ldots, 1]\}$ for R^n contains n elements. Therefore, $\dim R^n = n$.
(ii) The basis $\{1, x, x^2, \ldots, x^n\}$ for P_n contains $(n + 1)$ elements. Therefore, $\dim P_n = n + 1$.
(iii) The basis $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ for M_{22} contains 4 elements. Therefore, dim $M_{22} = 4$. In general $\dim M_{mn} = mn$.

EXAMPLE 2.58 Determine the dimension of set $S = \{[a, b, c, d] \mid c = a - b, d = a + b \text{ and } a, b, c, d \in R\}$

Solution: Here $S = \{(a, b, c, d) \mid c = a - b, d = a + b \text{ and } a, b, c, d \in R\}$

$$= \{[a, b, a - b, a + b] \mid a, b \in R\}$$

$$= \{[a, 0, a, a] + [0, b, -b, b] \mid a, b \in R\}$$

$$= \{a[1, 0, 1, 1] + b[0, 1, -1, 1] \mid a, b \in R\}$$

Thus the set S is generated by the two vectors $\mathbf{v}_1 = [1, 0, 1, 1]$ and $\mathbf{v}_2 = [0, 1, -1, 1]$.

That is, all the vectors of S are the linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Let $B = \{\mathbf{v}_1, \mathbf{v}_2\}$. Then S = span B.

If we can show that B is a linearly independent set, then B is a basis for S and dim S = number of vectors in basis B, i.e. dim $S = 2$.

Consider the vector equation for B

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2.$$

$$\alpha_1[1, 0, 1, 1] + \alpha_2[0, 1, -1, 1] = 0$$

$$\alpha_1 = 0, \ \ \alpha_2 = 0.$$

Thus the vector equation for B has the trivial solution. So B is a linearly independent set.

Hence dim $S = 2$.

EXAMPLE 2.59 Find the dimension of the set $S = \{[a, b, c, d] \mid a = b = c = d \text{ and } a, b, c, d \in R\}$.

Solution: Here $S = \{[a, b, c, d] \mid a = b = c = d \text{ and } a, b, c, d \in R\}$.

$$= \{[a, a, a, a] \mid a \in R\}$$

$$= \{a[1, 1, 1, 1] \mid a \in R\}.$$

Thus the set S is generated by the vector $\mathbf{v} = [1, 1, 1, 1]$. That is, all the vectors of S are the linear combination of V.

Let $B = \{\mathbf{v}\} = \{[1, 1, 1, 1]\}$. Then S = span B and also the set B is a linearly independent set, so B is a basis for S. Hence dim S = number of vectors in B, i.e. dim $S = 1$.

Some Important Properties

(i) The set S of any n linearly independent vectors of an n-dimensional vector space V is a basis of V.

(ii) If the set S of vectors of a finite-dimensional vector space V spans V, that is, V = span S, but is not a basis for V, then S can be reduced to a basis for V by removing the appropriate vector from S.

(iii) If W is a subspace of an n-dimensional vector space V, then dim $W \le n$. Moreover, dim $W = n$ if and only if $W = V$.

Theorem 2.16 [Extension Property]

Let S be a finite set of a finite-dimensional vector space V. If S is a linearly independent set but not a basis for V, then S can be extended to a basis for V by inserting an appropriate vector into S.

EXAMPLE 2.60 Find a standard basis vector that can be added to the set $\{[-1, 2, 3], [1, -2, -2]\}$ to produce a basis for R^3.

Solution: Let $\mathbf{v}_1 = [-1, 2, 3]$, $\mathbf{v}_2 = [1, -2, -2]$.

It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set. Also

$$[S] = \text{span } S = \{\alpha[-1, 2, 3] + \beta[1, -2, -2] \mid \alpha, \beta \in R\}$$

$$= \{[-\alpha + \beta, 2\alpha - 2\beta, 3\alpha - 2\beta) \mid \alpha, \beta \in R\}$$

Consider the standard basis vector $\mathbf{e}_1 = [1, 0, 0]$. Suppose

$$\mathbf{e}_1 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$$

$$[1, 0, 0] = \alpha[-1, 2, 3] + \beta[1, -2, -2]$$

$$1 = -\alpha + \beta,\ 0 = 2\alpha - 2\beta,\ 0 = 3\alpha - 2\beta$$

$\therefore$ $$\alpha = 0,\ \beta = 0$$

Thus $\mathbf{e}_1 = [1, 0, 0]$ cannot be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Therefore the set $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1\}$ is a linearly independent set of three vectors of R^3. Hence B is a basis for R^3.

EXAMPLE 2.61 Find the standard basis vector that can be added to the set $\{[2, 1, 1, 2], [-1, 1, 1, -1]\}$ to produce a basis for R^4.

Solution: Let $\mathbf{v}_1 = [2, 1, 1, 2]$, $\mathbf{v}_2 = [-1, 1, 1, -1]$.

It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set. Also

$$\begin{aligned}[S] &= \text{span } S = \{\alpha[2, 1, 1, 2] + \beta[-1, 1, 1, -1] \mid \alpha, \beta \in R\} \\ &= \{(2\alpha - \beta, \alpha + \beta, \alpha + \beta, 2\alpha - \beta) \mid \alpha, \beta \in R\}\end{aligned}$$

By inspection, the standard basis vector $\mathbf{e}_1 = [1, 0, 0, 0]$ does not belong to $[S]$, i.e., $\mathbf{e}_1 \notin [S]$.

Therefore the set $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1\}$ is a linearly independent set of R^4. Also,

$$\begin{aligned}[B_1] &= \{\alpha[2, 1, 1, 2] + \beta[-1, 1, 1, -1] + \gamma[1, 0, 0, 0] \mid \alpha, \beta, \gamma \in R\} \\ &= \{(2\alpha - \beta + \gamma, \alpha + \beta, \alpha + \beta, 2\alpha - \beta) \mid \alpha, \beta, \gamma \in R\}\end{aligned}$$

Again we can easily see that the standard basis vector $\mathbf{e}_2 = [0, 1, 0, 0]$ does not belong to $[B_1]$, i.e. $\mathbf{e}_2 \notin [B_1]$. Therefore the set $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2\}$ is a linearly independent set of four vectors of R^4. Therefore B_2 is a basis for R^4.

EXERCISE SET 4

1. Which of the following sets of vectors are bases for the indicated vector spaces.

(i) $\{[1, -1], [0, 6]\}$ for R^2

(ii) $\{[2, 3], [-4, -6]\}$ for R^2

(iii) $\{[1, 3, 4], [-1, 0, 2]\}$ for R^3

(iv) $\{[1, -1, 1], [1, 1, 2], [3, 0, -1]\}$ for R^3

(v) $\{1, t - 1, (t - 1)(t - 2)\}$ for P_2

2. (i) Find the co-ordinate vector of $\mathbf{v} = [4, -2, 3]$ relative to the basis $S = \{[2, 0, 0], [0, -4, 0], [-1, 0, 5]\}$.

(ii) Find the co-ordinate vector of $\mathbf{p} = 10 + 5x$ relative to the basis $S = \{1 - x + x^2, x + 2x^2, 3 - x^2\}$.

(iii) Find the co-ordinate vector of $M = \begin{bmatrix} -1 & 0 \\ 1 & -4 \end{bmatrix}$ relative to the basis

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} \right\}.$$

3. (i) Determine the dimension of set $S = \{[a, b, c, d] \mid a + c - d = 0 \text{ and } a, b, c, d \in R\}$.

(ii) Determine the dimension of set $S = \{[a, b, c, d] \mid a + 2b = 0 \text{ and } a, b, c, d \in R\}$.

4. Find a standard basis vector that can be added to the set $\{[1, -1, 0], [3, 1, -2]\}$ to produce a basis for R^3.

2.5 ROW SPACE, COLUMN SPACE AND NULL SPACE

Consider an $m \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

with entries in R. Then the rows of **A** can be described as vectors in R^n,

$$\begin{aligned} \mathbf{r}_1 &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix} \\ \mathbf{r}_2 &= \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix} \\ \vdots \quad & \quad \vdots \qquad\qquad \dots \quad \vdots \\ \mathbf{r}_m &= \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{aligned}$$

known as row vectors of **A**. Similarly, the columns of **A** can be described as vectors in R^m,

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

known as column vectors of **A**.

EXAMPLE 2.62 List the row vectors and column vectors of the matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 4 & 6 & -3 & 5 \\ 3 & 2 & 0 & -1 \end{bmatrix}$.

Solution: The row vectors of the given matrix **A** in R^4 are given by

$$\begin{aligned} \mathbf{r_1} &= \begin{bmatrix} 1 & -1 & 2 & 3 \end{bmatrix} \\ \mathbf{r_2} &= \begin{bmatrix} 4 & 6 & -3 & 5 \end{bmatrix} \\ \mathbf{r_3} &= \begin{bmatrix} 3 & 2 & 0 & -1 \end{bmatrix} \end{aligned}$$

The column vectors of the given matrix **A** in R^3 are given by

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}; \quad \mathbf{c}_2 = \begin{bmatrix} -1 \\ 6 \\ 2 \end{bmatrix}; \quad \mathbf{c}_3 = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}; \quad \mathbf{c}_4 = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$$

In this section, we will deal with three vector spaces that arise from the matrix **A**. These can be defined as follows.

Definition: Row Space, Column Space, Null Space

Suppose that **A** is an $m \times n$ matrix with entries in R. Then:

(i) The subspace span$\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ of R^n, where $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m$ are the row vectors of **A**, is called the *row space* of **A**.

(ii) The subspace span$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ of R^m, where $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the column vectors of **A**, is called the *column space* of **A**.

(iii) The solution space of the system of homogeneous linear equation $\mathbf{Ax} = \mathbf{0}$ is called the *null space* of **A**.

Remarks: Note that span $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m\}$ is a subspace of R^n, span$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is a subspace of R^m and also note that the null space of **A** is a subspace of R^n.

Row Space and Column Space

EXAMPLE 2.63 Find the row space and column space of the matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 5 \\ 2 & -3 & 6 \\ -1 & 2 & 0 \\ 4 & 1 & 1 \end{bmatrix}$.

Solution: The row vectors of the given matrix **A** are

$$\mathbf{r}_1 = [1 \quad -1 \quad 5]; \qquad \mathbf{r}_2 = [2 \quad -3 \quad 6]$$
$$\mathbf{r}_3 = [-1 \quad 2 \quad 0]; \qquad \mathbf{r}_4 = [4 \quad 1 \quad 1].$$

Therefore the row space of **A** is

$$\begin{aligned} \mathrm{R}(\mathbf{A}) &= \text{span } \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\} \\ &= \{\alpha_1\mathbf{r}_1 + \alpha_2\mathbf{r}_2 + \alpha_3\mathbf{r}_3 + \alpha_4\mathbf{r}_4 \mid \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in R\} \end{aligned}$$

Note that R(**A**) is a subspace of R^3.

The column vectors of **A** are

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 4 \end{bmatrix}; \qquad \mathbf{c}_2 = \begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \end{bmatrix}; \qquad \mathbf{c}_3 = \begin{bmatrix} 5 \\ 6 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore the column space of **A** is

$$\begin{aligned} \mathrm{C}(\mathbf{A}) &= \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\} \\ &= \{\alpha_1\mathbf{c}_1 + \alpha_2\mathbf{c}_2 + \alpha_3\mathbf{c}_3 \mid \alpha_1, \alpha_2, \alpha_3 \in R\} \end{aligned}$$

Note that C(**A**) is a subspace of R^4.

Now we will see some properties of row space and column space in the following theorem.

Theorem 2.17 [Properties of Row Space and Column Space]

Suppose that matrix **B** can be obtained from matrix **A** by elementary row operations. Then:

(i) The row space of **B** is identical to the row space of **A**.

(ii) Any collection of column vectors of **A** is linearly independent if and only if the corresponding collection of column vectors of **B** is linearly independent.

(iii) A set of column vectors of **A** forms a basis for the column space of **A** if and only if the corresponding collection of column vectors of **B** forms a basis for the column space of **B**.

The next theorem will be used to find the bases for row space and column space of a matrix.

Theorem 2.18 [Bases for Row Space and Column Space]
If a matrix **R** is in row-echelon form, then the row vectors with the leading 1s (the nonzero row vectors) form a basis for the row space of the matrix **R** and the column vectors with the leading 1s of the row vectors form a basis for the column space of the matrix **R**.

EXAMPLE 2.64 Find the bases for the row spaces and column spaces of the following matrices:

$$\text{(i)}\quad \mathbf{A}=\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 6\\ 0 & 0 & 0\end{bmatrix} \qquad \text{(ii)}\quad \mathbf{A}=\begin{bmatrix}1 & 4 & 5 & 2\\ 2 & 1 & 3 & 0\\ -1 & 3 & 2 & 2\end{bmatrix}$$

Solution:

(i) Here $\mathbf{A}=\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 6\\ 0 & 0 & 0\end{bmatrix}$

Since the given matrix **A** is in row-echelon form, by Theorem 2.18, the row vectors with the leading 1s form a basis for the row space of **A**. Hence

$$B_1=\{[1 \quad 0 \quad -2],\ [0 \quad 1 \quad 6]\}$$

is a basis for the row space of **A**.
The column vectors with the leading 1s of the row vectors form a basis for the column space of **A**. So

$$B_2=\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\ \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$$

is a basis for the column space of **A**.

(ii) Here the given matrix $\mathbf{A}=\begin{bmatrix}1 & 4 & 5 & 2\\ 2 & 1 & 3 & 0\\ -1 & 3 & 2 & 2\end{bmatrix}$ is not in row-echelon form. So first, we convert it into row-echelon form. By the elementary row operations, the row-echelon form of matrix **A** is

$$\mathbf{R}=\begin{bmatrix}1 & 4 & 5 & 2\\ 0 & 1 & 1 & 4/7\\ 0 & 0 & 0 & 0\end{bmatrix}$$

By Theorem 2.18, the nonzero row vectors of **R** form a basis for the row space of **R** and elementary row operations do not change the row space of the matrix. Therefore, those vectors also form a basis for the row space of **A**. Hence the basis for the row space of **A** is

$$B_1=\{[1 \quad 4 \quad 5 \quad 2], [0 \quad 1 \quad 1 \quad 4/7]\}$$

The first and second columns of **R** contain the leading 1s of the row vectors, so

$$\mathbf{c}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad \mathbf{c}_2' = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R. Thus the corresponding column vectors of **A**, namely,

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}; \qquad \mathbf{c}_2 = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

form a basis for the column space of **A**.

EXAMPLE 2.65 Find a basis for the row space of $\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$ consisting entirely of row vectors from **A**.

Solution: In Example 2.64, we saw that the basis vectors of row space **A** were not all row vectors of **A**, but the basis vectors of column space were all the column vectors of **A**. So to find a basis for the row space of **A** consisting entirely of row vectors from **A**, first we will find a basis for the column space of $\mathbf{A}^{\mathrm{T}}$ and then we will transpose again to convert column vectors back to row vectors.

By transposing **A**, we get

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 5 & 7 \\ -1 & -4 & -6 \\ 3 & -4 & 2 \end{bmatrix}$$

The row-echelon form of $\mathbf{A}^{\mathrm{T}}$ is

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 5 & 7 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The first and second columns of $\mathbf{R}_1$ contain the leading 1s of the row vectors. So

$$\mathbf{c}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad \mathbf{c}_2' = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R_1. Thus the corresponding column vectors of **A**, namely,

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}; \qquad \mathbf{c}_2 = \begin{bmatrix} 5 \\ -4 \\ -4 \end{bmatrix}$$

form a basis for the column space of $\mathbf{A}^{\mathrm{T}}$. So by again transposing $\mathbf{A}^{\mathrm{T}}$, we get

$$\mathbf{r}_1 = \begin{bmatrix} 1 & -1 & 3 \end{bmatrix}; \qquad \mathbf{r}_2 = \begin{bmatrix} 5 & -4 & -4 \end{bmatrix}$$

which form a basis for the row space of **A**.

EXAMPLE 2.66 Find a basis for the subspace of R^4 spanned by the given vectors $\mathbf{v}_1 = [1, 1, 0, 0]$, $\mathbf{v}_2 = [0, 0, 1, 1]$, $\mathbf{v}_3 = [-2, 0, 2, 2]$, $\mathbf{v}_4 = [0, -3, 0, 3]$.

Solution: Consider a matrix **A** with $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ as row vectors, that is,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -2 & 0 & 2 & 2 \\ 0 & -3 & 0 & 3 \end{bmatrix}$$

So the vector space V spanned by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ is same as the row space of **A**. Therefore, a basis for the row space of **A** is a basis for V. Reducing this matrix to row-echelon form, yields

$$\mathbf{R} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore, the nonzero row vectors

$$\mathbf{r}_1 = [1, 1, 0, 0]; \qquad \mathbf{r}_2 = [0, 1, 0, 0]$$

$$\mathbf{r}_3 = [0, 0, 1, 1]; \qquad \mathbf{r}_4 = [0, 0, 0, 1].$$

form a basis for the row space of **A** and so it is a basis for the subspace of R^n spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$.

EXAMPLE 2.67 Find a subset of the vectors that forms a basis for the space spanned by the vectors $\mathbf{v}_1 = [1, -2, 0, 3]$, $\mathbf{v}_2 = [2, -4, 0, 6)]$ $\mathbf{v}_3 = [-1, -1, 2, 0]$ and $\mathbf{v}_4 = [0, -1, 2, 3]$ and then express each vector as a linear combination of the basis vectors.

Solution: Consider a matrix **A** whose column vectors are $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$, that is,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix}$$

So a basis of the column space of **A** is a basis of the vector space V spanned by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$. Reducing this matrix to row-echelon form,

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first, third and fourth columns of $\mathbf{R}_1$ contain the leading 1s of the row vectors. Thus,

$$\mathbf{c}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{c}_3' = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{c}_4' = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of $\mathbf{R}_1$. Thus, the corresponding column vectors of $\mathbf{A}$, namely

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 3 \end{bmatrix}; \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 3 \end{bmatrix}$$

form a basis for the column space of $\mathbf{A}$. Hence $\mathbf{v}_1 = [1, -2, 0, 0]$, $\mathbf{v}_3 = [-1, 1, 2, 0]$ and $\mathbf{v}_4 = [0, -1, 2, 3]$ are basis vectors of space V.

The vector $\mathbf{c}_2'$ is not a basis vector. So it can be expressed as a linear combination of $\mathbf{c}_1'$, $\mathbf{c}_3'$ and $\mathbf{c}_4'$. By inspection, we have

$$\mathbf{c}_2' = 2\mathbf{c}_1'$$

The corresponding relationship is $\mathbf{v}_2 = 2\mathbf{v}_1$.

Null Space

Now we will discuss the solution spaces of the homogeneous equation $\mathbf{Ax} = \mathbf{0}$. Let us start with some examples.

EXAMPLE 2.68 Find the null space of matrix $\mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 3 & -1 & 0 \\ 2 & -4 & -5 \end{bmatrix}$.

Solution: The null space of $\mathbf{A}$ is the solution space of the homogeneous system

$$\mathbf{Ax} = \mathbf{0}.$$

$$\begin{aligned} -x_1 + x_2 + x_3 &= 0 \\ 3x_1 - x_2 &= 0 \\ 2x_1 - 4x_2 - 5x_3 &= 0 \end{aligned}$$

The augmented matrix for the system is

$$\begin{bmatrix} -1 & 1 & 1 & 0 \\ 3 & -1 & 0 & 0 \\ 2 & -4 & -5 & 0 \end{bmatrix}$$

Reducing this matrix to row-echelon form, we obtain

$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding system of equations is

$$2x_1 + x_3 = 0$$
$$2x_2 + 3x_3 = 0$$

Solving for the leading variables, we have

$$x_1 = -\frac{x_3}{2}; \qquad x_2 = -\frac{3}{2}x_3$$

Thus, the general solution is

$$x_1 = -\frac{s}{2}; \qquad x_2 = -\frac{3}{2}s; \qquad x_3 = s$$

This result can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}s \\ -\frac{3}{2}s \\ s \end{bmatrix} = s\begin{bmatrix} -\frac{1}{2} \\ -\frac{3}{2} \\ 1 \end{bmatrix}$$

Therefore the null space of **A** is $\mathrm{N}(\mathbf{A}) = \left\{ s\left[-\frac{1}{2}, -\frac{3}{2}, 1\right] \middle| s \in R \right\}$

Theorem 2.19 [Null Space]

Elementary row operations do not change the null space of a matrix.

EXAMPLE 2.69 Find a basis for the null space of a matrix $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$.

Solution: The null space of **A** is the solution space of the homogeneous system

$$\mathbf{Ax} = \mathbf{0}$$

$$\begin{aligned} x_1 + 4x_2 + 5x_3 + 6x_4 + 9x_5 &= 0 \\ 3x_1 - 2x_2 + x_3 + 4x_4 - x_5 &= 0 \\ -x_1 \qquad - x_3 - 2x_4 - x_5 &= 0 \\ 2x_1 + 3x_2 + 5x_3 + 7x_4 + 8x_5 &= 0 \end{aligned}$$

The augmented matrix for this system of linear equations is

$$\begin{bmatrix} 1 & 4 & 5 & 6 & 9 & 0 \\ 3 & -2 & 1 & 4 & -1 & 0 \\ -1 & 0 & -1 & -2 & -1 & 0 \\ 2 & 3 & 5 & 7 & 8 & 0 \end{bmatrix}$$

Reducing this matrix to row-echelon form, we obtain

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding system of linear equations is

$$x_1 + x_3 + 2x_4 + x_5 = 0$$
$$x_2 + x_3 + x_4 + 2x_5 = 0$$

Solving for the leading variables, we have

$$x_1 = -x_3 - 2x_4 - x_5; \qquad x_2 = -x_3 - x_4 - 2x_5$$

Thus the general solution is

$$\begin{aligned} x_1 &= -r - 2s - t \\ x_2 &= -r - s - 2t \\ x_3 &= r \\ x_4 &= s \\ x_5 &= t \end{aligned}$$

This result can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -r-2s-t \\ -r-s-2t \\ r \\ s \\ t \end{bmatrix} = r\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore the null space of **A** is

$$\mathrm{N}(\mathbf{A}) = \{r[-1, -1, 1, 0, 0], s[-2, -1, 0, 1, 0], t[-1, -2, 0, 0, 1]\}$$

and a basis **B** is $\{[-1, -1, 1, 0, 0], [-1, -1, 0, 1, 0], [-1, -2, 0, 0, 1]\}$

Theorem 2.20 [Column Space]

A system of linear equations $\mathbf{Ax} = \mathbf{b}$ is consistent if and only if **b** is in the column space of **A**.

EXAMPLE 2.70 Determine whether **b** is in the column space of **A**, and if so express **b** as a linear combination of the column vectors of **A**.

(i) $\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$

Solution:

(i) The given system of equations is

$$\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$$

The augmented matrix of the system is

$$\begin{bmatrix} 2 & 3 & 8 \\ -1 & 4 & 7 \end{bmatrix}$$

By Gauss-elimination method,

$$\begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

$$x_1 = 1; \qquad x_2 = 2$$

Therefore the system is consistent and so **b** is in the column space of **A**. Moreover,

$$\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$$

i.e.
$$1\begin{bmatrix} 2 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$$

Thus $\mathbf{b} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$ can be expressed as a linear combination of column vectors $\mathbf{c}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$; $\mathbf{c}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

(ii) The given system of equations is

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

The augmented matrix of the system is

$$\begin{bmatrix} 1 & 1 & 2 & -1 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 2 \end{bmatrix}$$

By Gauss-elimination method,

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The last row suggests that the system is inconsistent. Therefore, **b** is not in the column space of **A**.

(iii) The given system of equations is

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$$

The augmented matrix of the system is

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 4 \\ 0 & 1 & 2 & 1 & 3 \\ 1 & 2 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 & 7 \end{bmatrix}$$

By Gauss-elimination method,

$$x_1 = 30; \qquad x_2 = -15; \qquad x_3 = 7; \qquad x_4 = 4$$

Therefore the system is consistent and so **b** is in the column space of **A**. Moreover,

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 30 \\ -15 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$$

i.e.
$$30\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - 15\begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix} + 7\begin{bmatrix} 0 \\ 2 \\ -1 \\ 2 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$$

Thus $\mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$ can be expressed as a linear combination of column vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \qquad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{c}_4 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 2 \end{bmatrix}.$$

Theorem 2.21 [Nonhomogeneous System]

Suppose that $\mathbf{A}$ is a matrix with entries in R. Suppose further that $\mathbf{x}_0$ is a solution of the nonhomogeneous system of equations $\mathbf{Ax} = \mathbf{b}$, and that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a basis for the null space of $\mathbf{A}$. Then every solution of the system $\mathbf{Ax} = \mathbf{b}$ can be written in the form

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{c}_1\mathbf{v}_1 + \mathbf{c}_2\mathbf{v}_2 + \cdots + \mathbf{c}_r\mathbf{v}_r \tag{2.1}$$

where $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r \in R$.

On the other hand, every vector of the form (2.1) is a solution to the system $\mathbf{Ax} = \mathbf{b}$.

***Geometrical Interpretation of Theorem* 2.21**

Theorem 2.21 has a nice geometrical interpretation in R^2. The solution space of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ in R^2 is the origin or a line passing through origin or all of R^2 (Figure 2.8). From Theorem 2.21, the solution of the nonhomogeneous system $\mathbf{Ax} = \mathbf{b}$ can be obtained by adding any fixed solution $\mathbf{x}_0$ to the solutions of $\mathbf{Ax} = \mathbf{0}$.

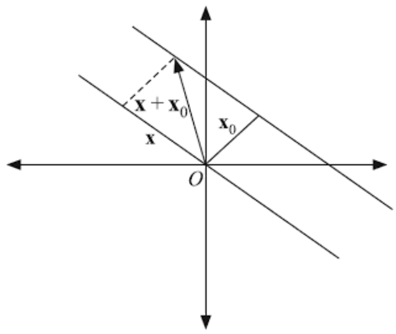


Figure 2.8

EXAMPLE 2.71 Find the vector form of the general solution of the system

$$x_1 - 3x_2 = 1$$
$$2x_1 - 6x_2 = 2$$

and then use that result to find the vector form of the general solution of the homogeneous system.

Solution: The given system of linear equation is

$$x_1 - 3x_2 = 1$$
$$2x_1 - 6x_2 = 2$$

The augmented matrix of this system is

$$\begin{bmatrix} 1 & -3 & 1 \\ 2 & -6 & 2 \end{bmatrix}$$

By Gauss-elimination method,

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$x_1 - 3x_2 = 1$$

$$x_1 = 1 + 3s; \qquad x_2 = s$$

The vector form of this result is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1+3s \\ s \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

which is a general solution of the given nonhomogeneous system where $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a particular solution.

From Theorem 2.21, the general solution of the homogeneous system is

$$x_1 - 3x_2 = 0$$
$$2x_1 - 6x_2 = 0$$

is
$$\mathbf{x} = s\begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

EXAMPLE 2.72 Find the vector form of the general solution of

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 4 \\ -2x_1 + x_2 + 2x_3 + x_4 &= -1 \\ -x_1 + 3x_2 - x_3 + 2x_4 &= 3 \\ 4x_1 - 7x_2 \qquad - 5x_4 &= -5 \end{aligned}$$

and then use that result to find the vector form of the general solution of $\mathbf{Ax} = \mathbf{0}$.

Solution: The augmented matrix of the given system of equation is

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 4 \\ -2 & 1 & 2 & 1 & -1 \\ -1 & 3 & -1 & 2 & 3 \\ 4 & -7 & 0 & -5 & -5 \end{bmatrix}$$

Using the Gauss-elimination method, we obtain

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 4 \\ 0 & 5 & -4 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding equations are

$$x_1 + 2x_2 - 3x_3 + x_4 = 4$$
$$5x_2 - 4x_3 + 3x_4 = 7$$

Therefore the general solution of the given system is

$$x_1 = \frac{6}{5} + \frac{7}{5}r + \frac{1}{5}s$$

$$x_2 = \frac{7}{5} + \frac{4}{5}r - \frac{3}{5}s$$
$$x_3 = r$$
$$x_4 = s$$

The vector form of this solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{6}{5} + \frac{7}{5}r + \frac{1}{5}s \\ \frac{7}{5} + \frac{4}{5}r - \frac{3}{5}s \\ r \\ s \end{bmatrix} = \begin{bmatrix} \frac{6}{5} \\ \frac{7}{5} \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{1}{5} \\ -\frac{3}{5} \\ 0 \\ 1 \end{bmatrix}$$

Here $\mathbf{x}_0 = \begin{bmatrix} \frac{6}{5} \\ \frac{7}{5} \\ 0 \\ 0 \end{bmatrix}$ is a particular solution of the given non-homogeneous system. From Theorem 2.21, the general solution of the homogeneous system

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 0 \\ -2x_1 + x_2 + 2x_3 + x_4 &= 0 \\ -x_1 + 3x_2 - x_3 + 2x_4 &= 0 \\ 4x_1 - 7x_2 - 5x_4 &= 0 \end{aligned}$$

is

$$\mathbf{x} = r\begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{1}{5} \\ -\frac{3}{5} \\ 0 \\ 1 \end{bmatrix}$$

EXERCISE SET 5

1. List the row vectors and column vectors of the following matrices:

(i) $\begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 2 & 3 \\ -5 & 6 & 1 \\ 3 & 2 & -4 \end{bmatrix}$ (iii) $\begin{bmatrix} -1 & 2 & 6 & 0 \\ 3 & -7 & -8 & 9 \end{bmatrix}$

2. Find the row space and column space of each matrix of Exercise 1.

3. Find the bases for the row space and the column space of $\mathbf{A} = \begin{bmatrix} 2 & -4 & 1 & 2 & -2 & -3 \\ -1 & 2 & 0 & 0 & 1 & -1 \\ 10 & -4 & -2 & 4 & -2 & 4 \end{bmatrix}$.

4. Find a basis for the row space of $\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 & 3 \\ 2 & -5 & -3 & 6 \\ 0 & 1 & 3 & 0 \\ 2 & -1 & 4 & -7 \\ 5 & -8 & 1 & 2 \end{bmatrix}$ consisting entirely of row vectors from **A**.

5. Find a basis for the subspace of R^5 spanned by the given vectors $\mathbf{v}_1 = [-1, 2, -1, 5, 6]$, $\mathbf{v}_2 = [4, -4, -4, -12, -8]$, $\mathbf{v}_3 = [2, 0, -6, -2, 4]$, $\mathbf{v}_4 = [-3, 1, 7, -2, 12]$.

6. Find a subset of the vectors that forms a basis for the space spanned by the vectors

$$\mathbf{v}_1 = [1, -2, 0, 0, 3],\ \mathbf{v}_2 = [2, -5, -3, -2, 6],\ \mathbf{v}_3 = [0, 5, 15, 10, 0],\ \mathbf{v}_4 = [2, 6, 18, 8, 6].$$

7. Find the null space of matrix $\mathbf{A} = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$ generated by basis vectors.

8. Find a basis for the null space of matrix $\mathbf{A} = \begin{bmatrix} 2 & -4 & 1 & 2 & -2 & -3 \\ -1 & 2 & 0 & 0 & 1 & -1 \\ 10 & -4 & -2 & 4 & -2 & 4 \end{bmatrix}$.

9. Determine whether **b** is in the column space of **A** and if so express **b** as a linear combination of the column vectors of **A**.

(i) $\mathbf{A} = \begin{bmatrix} -2 & 1 & -1 \\ 1 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4 \\ 13 \\ -1 \end{bmatrix}$ (ii) $\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ -1 & 1 & -2 \\ 2 & -1 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$

10. Find the vector form of the general solution of the system

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 5 \\ x_1 \qquad + x_3 &= -2 \\ 2x_1 + x_2 + 3x_3 &= 3 \end{aligned}$$

and then use that result to find the vector form of the general solution of the homogeneous system.

2.6 RANK AND NULLITY

In the previous section, we defined row space and null space of a matrix and calculated their bases. By continuing this discussion, we will here concentrate on the dimensions of row space and null space and see some nice relationships between them.

If we carefully observe the results of solved examples of the previous section, we can then come out with the following result.

Theorem 2.22 [Dimensions of Row Space and Null Space]

For any matrix **A** with entries in R, the dimension of the row space is the same as the dimension of the column space.

From the result of Theorem 2.22, we have the following definition.

Definition: Rank of a Matrix

The common value of the dimension of row space and column space of a matrix **A** is called the rank of a matrix **A** and it is denoted by

$$\text{rank } \mathbf{A}.$$

Definition: Nullity of a Matrix

The dimension of the null space of a matrix is called the nullity of **A** and denoted by

$$\text{nullity } \mathbf{A}$$

Theorem 2.23 [Rank of a Matrix]

If **A** is any matrix, then

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^{\mathrm{T}}.$$

EXAMPLE 2.73 Find the rank and nullity of the matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$.

Solution: By using the elementary row operations, the reduced row-echelon form of **A** is

$$\begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -19 \\ 0 & 0 & 0 \end{bmatrix}$$

From Theorem 2.18, the basis vectors of the row space of matrix **A** are $\mathbf{r}_1 = [1, -1, 3]$ and $\mathbf{r}_2 = [0, 1, -19]$. So the dimension of the row space is 2.

That is, $$\text{rank } \mathbf{A} = 2$$

The corresponding system of equations is

$$\begin{aligned} x_1 - x_2 + 3x_3 &= 0 \\ x_2 - 19x_3 &= 0 \end{aligned}$$

Therefore the general solution of this system is

$$x_1 = 16s: \qquad x_2 = 19s; \qquad x_3 = s$$

The vector form of the general solution is $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = s\begin{bmatrix} 16 \\ 19 \\ 1 \end{bmatrix}$.

So the basis vector of null space is [16, 19, 1]. Thus the dimension of null space is 1, that is

$$\text{nullity } \mathbf{A} = 1.$$

EXAMPLE 2.74 Verify the result

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^{\mathrm{T}}$$

for matrix $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$.

Solution: The reduced row-echelon form of **A** is

$$\begin{bmatrix} 1 & 4 & 5 & 2 \\ 0 & 1 & 1 & 4/7 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The basis vectors of the row space of **A** are $\mathbf{r}_1 = [1, 4, 5, 2]$ and $\mathbf{r}_2 = [0, 1, 1, 4/7]$. Thus the dimension of the row space of **A** is 2.

That is, $$\text{rank } \mathbf{A} = 2$$

The basis for the row space of $\mathbf{A}^{\mathrm{T}}$ is the same as the basis for the column space of **A** and the basis vectors of the column space of **A** are $\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\mathbf{c}_2 = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$.

So $$\text{rank } \mathbf{A}^{\mathrm{T}} = \text{rank } C(\mathbf{A}) = 2$$

Hence $$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^{\mathrm{T}}$$

Theorem 2.24 [Dimension Theorem]

For any matrix **A** with entries in R, the sum of the dimension of its column space and the dimension of its null space is equal to the number of columns of **A**. That is,

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = n \qquad \text{where } n \text{ is the number of columns of matrix } \mathbf{A}.$$

EXAMPLE 2.75 Verify the dimension theorem for the following matrices:

(i) $\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$ (ii) $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$

Solution:

(i) See Example 2.73,

rank **A** = 2 nullity **A** = 1 number of columns = n = 3

rank **A** + nullity **A** = 2 + 1 = 3 = n

Hence the dimension theorem is verified.

(ii) For the given matrix $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$, the reduced row-echelon form is

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{i}$$

From Theorem 2.18, the basis vectors of the row space of **A** are $\mathbf{r}_1$ = [1, 0, 1, 2, 1], $\mathbf{r}_2$ = [(0, 1, 1, 1, 2].

So the dimension of the row space of **A** is 2.

That is, rank **A** = 2.

The corresponding system of equations of (i) is

$$\begin{aligned} x_1 \quad + x_3 + 2x_4 + \ x_5 &= 0 \\ x_2 + x_3 + \ x_4 + 2x_5 &= 0 \end{aligned}$$

Therefore the general solution of the system is

$$x_1 = -r - 2s - t; \qquad x_2 = -r - s - 2t; \qquad x_3 = r; \qquad x_4 = s; \qquad x_5 = t.$$

The vector form of the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -r-2s-t \\ -r-s-2t \\ r \\ s \\ t \end{bmatrix} = r\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The basis vectors for the null space of a matrix **A** are

$$\mathbf{c}'_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{c}'_2 = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{c}'_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, the dimension of the null space of **A** is 3.

That is, nullity $\mathbf{A} = 3$.

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = 2 + 3 = 5 = n$$

Since the number of columns in **A** is $n = 5$, the dimension theorem is verified.

Remark: **[Maximum Value for Rank]**

(i) Let **A** be an $m \times n$ matrix. Then the possible maximum value of rank of **A** is the minimum of m and n. That is,

$$\text{rank } \mathbf{A} \le \min (m, n).$$

(ii) If **A** is an $m \times n$ matrix and **B** is an $n \times p$ matrix, then

$$\text{rank } (\mathbf{AB}) \le \min (\text{rank } \mathbf{A}, \text{rank } \mathbf{B}).$$

EXAMPLE 2.76 Suppose **A** be a 4×6 matrix. Find the maximum value of rank of **A**.

Solution: Since the matrix **A** has order 4×6, the maximum value of rank of **A** is minimum of 4 and 6. That is,

$$\text{maximum of rank } \mathbf{A} = \min (4, 6) = 4$$

EXERCISE SET 6

1. Find the rank and nullity of matrix $\mathbf{A} = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 1 & 4 & -7 & 3 & -2 \\ 1 & 5 & -9 & 5 & -9 \\ 0 & 3 & -6 & 2 & -1 \end{bmatrix}$.

2. Find the rank and nullity of matrix $\mathbf{B} = \begin{bmatrix} 2 & -4 & 1 & 2 & -2 & -3 \\ -1 & 2 & 0 & 0 & 1 & -1 \\ 10 & -4 & -2 & 4 & -2 & 4 \end{bmatrix}$.

3. Verify the result

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^{\mathrm{T}}$$

for matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & -8 \\ 4 & -3 & -7 \\ 1 & 12 & -3 \end{bmatrix}$.

4. Verify the dimension theorem for the following matrices:

(i) $\begin{bmatrix} 5 & 9 & 3 \\ 3 & -5 & -6 \\ 1 & 5 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 4 & -1 & 5 \\ 1 & 2 & 3 & -1 & 3 \\ 1 & 2 & 0 & -4 & -3 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 5 \\ 3 & 4 & 11 & 2 \end{bmatrix}$

5. Suppose **A** is a 5×3 matrix and **B** a 3×5 matrix. Find the maximum values of ranks of **A** and **B**.

SUMMARY

Vector A vector is a quantity that has both magnitude and direction.

Vector Addition If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ are two vectors in R^n, then the addition of $\mathbf{u}$ and $\mathbf{v}$ is defined as: $(\mathbf{u} + \mathbf{v}) = [u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots, u_n + v_n]$

Scalar Multiplication If $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ is a vector and k is a nonzero scalar (i.e. k is any number), then the scalar multiple, $k\mathbf{v}$, is defined as: $k\mathbf{v} = [kv_1, kv_2, kv_3, \ldots, kv_n]$.

Subtraction If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$, then the difference $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = [u_1 - v_1, u_2 - v_2, u_3 - v_3, \ldots, u_n - v_n]$.

Theorem [Properties of Vector in R^n] If $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors in R^n and k and m are scalars, then

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
(ii) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
(iv) $\mathbf{u} - \mathbf{u} = \mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
(v) $1\mathbf{u} = \mathbf{u}$
(vi) $(km)\mathbf{u} = k(m\mathbf{u}) = m(k\mathbf{u})$
(vii) $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$
(viii) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

Norm If $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ is a vector in R^n, then the magnitude of vector is called the norm of the vector $\mathbf{v}$ and it is denoted by $\|\mathbf{v}\|$ and defined by the formula, $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

Theorem [Properties of Norm in R^n] Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in R^n and k be any scalar, then

(i) $\|\mathbf{u}\| \geq 0$ and $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
(ii) $\|k\mathbf{u}\| = |k|\,\|\mathbf{u}\|$
(iii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Distance The distance between two vectors, $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ in R^n is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} + \mathbf{v}\| = \sqrt{(u_1 + v_1)^2 + (u_2 + v_2)^2 + \cdots + (u_n + v_n)^2}$$

Theorem [Properties of Distance in R^n] Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be vectors in R^n and k be any scalar, then

(i) $d(\mathbf{u}, \mathbf{v}) \geq 0$ and $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$.
(ii) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
(iii) $d(\mathbf{u} + \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

Dot Product in R^n If $\mathbf{u}$ and $\mathbf{v}$ are two nonzero vectors in R^n and θ is the angle between them, then the dot product or inner product $\mathbf{u} \cdot \mathbf{v}$ in R^n is the real number defined by

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

Another Definition: Dot Product in R^n If $\mathbf{u} = [u_1, u_2, u_3, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, v_3, \ldots, v_n]$ are any vectors in R^n, then the dot product or inner product $\mathbf{u} \cdot \mathbf{v}$ is defined by the formula,

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Theorem [Properties of the Dot Product] Let **u**, **v**, **w** be vectors in R^3 (or in R^2) and α, β be real numbers. Then

(i) $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$. Hence $\mathbf{u} \cdot \mathbf{u} \geq 0$

(ii) $\mathbf{u} \cdot \mathbf{u} = 0$ if and if $\mathbf{u} = \mathbf{0}$.

(iii) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

(iv) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

(v) $(\alpha\mathbf{u}) \cdot \mathbf{v} = \alpha(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\alpha\mathbf{v})$

(vi) If θ is an angle between **u** and **v**, then

$$\theta \text{ is acute if and only if } \mathbf{u} \cdot \mathbf{v} > 0$$

$$\theta \text{ is obtuse if and only if } \mathbf{u} \cdot \mathbf{v} < 0$$

$$\theta = \frac{\pi}{2} \text{ if and only if } \mathbf{u} \cdot \mathbf{v} = 0.$$

Orthogonal Vectors Two vectors **u** and **v** in R^n are called orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

Theorem [Cauchy–Schwarz Inequality in R^n] If **u** and **v** are vectors in R^n, then $|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}||\ ||\mathbf{v}||$.

Theorem [Pythagorus Theorem] If **u** and **v** are orthogonal vectors in R^n, then

$$||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$$

The Dot Product in Form of Matrix Multiplication The dot product of the vectors **u** and **v** in R^n can be expressed in the form of matrix operation as

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{u}.$$

Properties If **A** is an $n \times n$ matrix, then

(i) $\mathbf{Au} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{A}^{\mathrm{T}}\mathbf{v}$ (ii) $\mathbf{u} \cdot \mathbf{Av} = \mathbf{A}^{\mathrm{T}}\mathbf{u} \cdot \mathbf{v}$

Vector Space If V is a non-empty set of objects called vectors with two operations: (i) addition ($\mathbf{u} + \mathbf{v}$, $\mathbf{u}, \mathbf{v} \in V$) (ii) scalar multiplication ($c\mathbf{u}$, c is real scalar) and satisfies the following properties, then V is called vector space over R or real vector space.

(i) *Closure for addition:* For every $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} \in V$

(ii) *Commutative for addition:* For every $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(iii) *Associative law for addition:* For every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

(iv) *Zero vector:* There exists a unique vector in V, called zero vector and denoted by **0**, such that $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

(v) *Additive inverse:* For every **u** in V there exists a unique vector $-\mathbf{u}$ in V called negative of **u** or additive inverse of **u**, such that

$$\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$$

(vi) *Closure property for scalar multiplication:* For every scalar $k \in R$ and $\mathbf{u} \in V$, $k\mathbf{u} \in V$

(vii) *Distributive law for addition:* For every scalar $k \in R$ and $\mathbf{u}, \mathbf{v} \in V$, $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

(viii) *Distributive law for scalar multiplication:* For every scalar $k, m \in R$ and $\mathbf{u} \in V$

$$(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$$

(ix) *Associative law for scalar multiplication:* For every scalar k, $m \in R$ and $\mathbf{u} \in V$, $k(m\mathbf{u}) = (km)\mathbf{u}$

(x) For every $\mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$.

Linear Combination If a vector $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n,$$

then the vector $\mathbf{w}$ is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars.

Span Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a subset of a vector space V. Then the set of all linear combinations of the vectors in S is called the span of S. It is denoted by span (S) and it can also be represented as

$$\begin{aligned}\text{span}\,(S) &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}\\ &= \{\mathbf{w} \mid \mathbf{w} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_r\mathbf{v}_r, \text{ each } \alpha_i \text{ is scalar, } 1 \le i \le r\}\end{aligned}$$

Theorem If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are vectors in a vector space V, then

(i) span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is subspace of V.

(ii) span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is the smallest subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, that is, every other subspace of V that contains $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ must contain span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$.

Theorem If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $S_1 = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are two sets of vectors in a space V, then

$$\text{span}\,\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$$

if and only if each vector in S is a linear combination of those in S_1 and each vector in S_1 is a linear combination of those in S_1.

Linearly Independent and Linearly Dependent A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of vectors is said to be *linearly independent*, if the vector equation

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$$

has only one trivial solution $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. If this vector equation has other non-trivial solutions, that is, $\alpha_1, \alpha_2, \ldots, \alpha_n$ not all zero, then S is called a *linearly dependent* set.

Theorem A set S with two or more vectors is

(i) linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.

(ii) linearly independent if and only if no vector in S is expressible as a linear combination of the other vector in S.

Corollary 1 A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.

Corollary 2 A finite set of vectors that contains the zero vector is linearly dependent.

Recall

(i) A homogeneous system $\mathbf{Ax} = \mathbf{0}$ of n linear equations in n unknowns

(ii) A homogeneous system $\mathbf{Ax} = \mathbf{0}$ of n linear equations in m unknowns has the non-trivial solution if $m > n$.

Theorem [Linear Dependence] If the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of vectors of R^n (or P_n) and $m > n$, then S is linearly dependent.

Linear Independence of Functions Let $f_1(x), f_2(x), \ldots, f_n(x)$ be $(n-1)$ times differentiable functions of $F(-\infty, \infty)$. Then the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1^1(x) & f_2^1(x) & \cdots & f_n^1(x) \\ \cdot & \cdot & \cdots & \cdot \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \cdots & f_n^{n-1}(x) \end{vmatrix}$$

is called Wronskian of $f_1(x), f_2(x), \ldots, f_n(x)$.

Theorem [Linear Dependence of Vectors] If the functions $f_1, f_2, \ldots, f_n$ have continuous derivatives on the interval $(-\infty, \infty)$ and if the Wronskian $W(\mathrm{x})$ of these functions is not identically zero on $(-\infty, \infty)$, then these functions form a linearly independent set of vectors in C^{n-1} $(-\infty, \infty)$.

Basis Suppose $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors from the space V. Then S is called a basis for V if both the following conditions hold.

(i) S spans the vector space V, that is, $V = \text{span}\,(S)$.
(ii) S is a linearly independent set of vectors.

Theorem [Basis] Suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for vector space V. Then every vector $\mathbf{u} \in V$ can be expressed uniquely in the form

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \qquad \text{where } c_1, c_2, \ldots, c_n \text{ are scalars.}$$

Co-ordinates Relatives to a Basis If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, and

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \ldots, c_n$ are called the co-ordinates of $\mathbf{v}$ relative to the basis S. It is denoted by $(\mathbf{v})_S = [c_1, c_2, \ldots, c_n]$.

Finite-Dimensional and Infinite-Dimensional Vector Spaces A vector spaces V over R is said to be *finite-dimensional* if it has a basis containing only finitely many elements, otherwise it is called an *infinite-dimensional* vector space.

Theorem [Finite Dimensional Vector Space] Let V be a finite-dimensional vector space and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis for V. Then:

(i) If the set has more than n vectors, it is linearly dependent.
(ii) If the set has fewer than n vectors, it does not span V.

Theorem [Basis for a Vector Space] Any two bases for a vector space must contain the same number of elements. This theorem suggests the following definition.

Definition: Dimension of Vector Space The number of vectors in a basis for a finite-dimensional vector space V is called the dimension of V. It is denoted by dim V.

Some Important Properties

(i) The set S of any n linearly independent vectors of an n-dimensional vector space V is a basis of V.
(ii) If the set S of vectors of a finite-dimensional vector space V spans V, that is , $V = \text{span } S$, but is not a basis for V, then S can be reduced to a basis for V by removing the appropriate vector from S.
(iii) If W is a subspace of an n-dimensional vector space V, then dim $W \le n$. Moreover, dim $W = n$ if and only if $W = V$.

Theorem [Extension Property] Let S be a finite set of a finite-dimensional vector space V. If S is a linearly independent set but not a basis for V, then S can be extended to a basis for V by inserting an appropriate vector into S.

Row Space, Column Space and Null Space Suppose that $\mathbf{A}$ is an $m \times n$ matrix with entries in R. Then

(i) The subspace span$\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m\}$ of R^n, where $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ are the row vectors of $\mathbf{A}$, is called the **row space** of $\mathbf{A}$.

(ii) The subspace span$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ of R^m, where $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are the column vectors of $\mathbf{A}$, is called the **column space** of $\mathbf{A}$.

(iii) The solution space of the system of homogeneous linear equation $\mathbf{A}x = \mathbf{0}$ is called the **null space** of $\mathbf{A}$.

Remarks Note that span $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_m\}$ is a subspace of R^n, span $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is a subspace of R^m and also note that the null space of $\mathbf{A}$ is a subspace of R^n.

Theorem [Properties of Row Space and Column Space] Suppose that matrix $\mathbf{B}$ can be obtained from matrix $\mathbf{A}$ by elementary row operations. Then:

(i) The row space of $\mathbf{B}$ is identical to the row space of $\mathbf{A}$.

(ii) Any collection of column vectors of $\mathbf{A}$ is linearly independent if and only if the corresponding collection of column vectors of $\mathbf{B}$ is linearly independent.

(iii) A set of column vectors of $\mathbf{A}$ forms a basis for the column space of $\mathbf{A}$ if and only if the corresponding collection of column vectors of $\mathbf{B}$ forms a basis for the column space of $\mathbf{B}$.

Theorem [Bases for Row Space and Column Space] If a matrix $\mathbf{R}$ is in row-echelon form, then the row vectors with the leading 1s (the nonzero row vectors) form a basis for the row space of the matrix $\mathbf{R}$ and the column vectors with the leading 1s of the row vectors form a basis for the column space of the matrix $\mathbf{R}$.

Theorem [Null Space] Elementary row operations do not change the null space of matrix.

Theorem [Column Space] A system of linear equation $\mathbf{Ax} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of $\mathbf{A}$.

Theorem [Nonhomogeneous System] Suppose that $\mathbf{A}$ is a matrix with entries in R. Suppose further that $\mathbf{x}_0$ is a solution of the nonhomogeneous system of equation $\mathbf{Ax} = \mathbf{b}$, and that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a basis for the null space of $\mathbf{A}$. Then every solution of the system $\mathbf{Ax} = \mathbf{b}$ can be written in the form

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{c}_1\mathbf{v}_1 + \mathbf{c}_2\mathbf{v}_2 + \cdots + \mathbf{c}_r\mathbf{v}_r$$

where $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r \in R$.

Theorem [Dimensions of Row Space and Null Space] For any matrix $\mathbf{A}$ with entries in R, the dimension of the row space is the same as the dimension of the column space.

Rank of a Matrix The common value of the dimension of row space and column space of a matrix $\mathbf{A}$ is called the rank of a matrix $\mathbf{A}$ and it is denoted by

$$\text{rank } \mathbf{A}.$$

Nullity of a Matrix The dimension of the null space of a matrix is called the nullity of $\mathbf{A}$ and denoted by

$$\text{nullity } \mathbf{A}.$$

Theorem [Rank of a Matrix] If $\mathbf{A}$ is any matrix, then

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^{\mathrm{T}}.$$

Theorem [Dimension Theorem] For any matrix $\mathbf{A}$ with entries in R, the sum of the dimension of its column space and the dimension of its null space is equal to the number of columns of $\mathbf{A}$. That is,

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = n \qquad \text{where } n \text{ is number of columns of matrix } \mathbf{A}.$$

Remark **[Maximum Value for Rank]**

(i) Let $\mathbf{A}$ be an $m \times n$ matrix. Then the possible maximum value of rank of $\mathbf{A}$ is the minimum of m and n. That is,

$$\text{rank } \mathbf{A} \le \min(m, n).$$

(ii) If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix, then

$$\text{rank } \mathbf{AB} \le \min(\text{rank } \mathbf{A}, \text{rank } \mathbf{B}).$$

3 Linear Transformations

3.1 INTRODUCTION

In the previous chapter, we discussed the vector space and learnt about its properties. But in the application of vector space, we need to associate the elements of one vector space to the elements of the other vector space by defining a function between them. Such a function is called *transformation*, *mapping* or *operator*. In this chapter, we will discuss the simplest class of such functions which is called *linear transformation*. The whole discussion about the linear transformation will be developed in two forms.

(i) The linear transformation from a general vector space V to W.

(ii) The linear transformation from an Euclidean vector space R^n to R^m.

First we introduce some notations and terminology which will be used in this chapter.

Let V and W be two non-empty sets. Then the function whose domain is V and co-domain is W is denoted by

$$T: V \to W$$

The vector $\mathbf{w}$ in W assigned to a vector $\mathbf{v}$ in V is called the *image* of $\mathbf{v}$ and is denoted by $T(\mathbf{v})$. The set of all images of all vectors of the domain V is called the *range* of T and it is denoted by $T(V)$ which is a subset of the co-domain W.

3.2 LINEAR TRANSFORMATION: DEFINITION AND EXAMPLES

In this section we will define the linear transformation and go through some examples.

Definition: Linear Transformation

Let V and W be two vector spaces with the same sets of scalars. Then a function, as defined above, $T: V \to W$ is called a linear transformation of V into W if it satisfies the following properties:

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \qquad \text{for all } \mathbf{u} \text{ and } \mathbf{v} \text{ in } V \tag{3.1}$$

$$T(\alpha\mathbf{v}) = \alpha T(\mathbf{v}) \qquad \text{for every } \mathbf{v} \text{ in } V \text{ and every scalar } \alpha \tag{3.2}$$

Remark: The above mentioned two properties can be combined into a single property as follows.

$$T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v}), \qquad \text{for all } \mathbf{u} \text{ and } \mathbf{v} \text{ in } V \text{ and all scalars } \alpha \text{ and } \beta \tag{3.3}$$

That is, to check the linearity of a transformation , it is enough to show that it satisfies the property (3.3). If the vector spaces V and W are Euclidean spaces R^n and R^m, then the above definition of linear transformation is reduced to the following form.

Definition: Linear Transformation in Euclidean Vector Spaces

The function $T: R^n \to R^m$ is called linear transformation if

$$T(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_m)$$

where

$$\begin{aligned}
u_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
u_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \qquad \vdots \qquad \cdots \quad \vdots \\
u_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \qquad 1 \le i \le m, \quad 1 \le j \le n
\end{aligned}$$

Remark: If $V = W$, then the linear transformation $T: V \to V$ is called a *linear operator* on V.

EXAMPLE 3.1 Check the linearity of the following transformation

$$T: R^3 \to R^2$$

defined by $T(x_1, x_2, x_3) = (x_1 + 2x_3, x_1 - x_2)$.

Solution: Let $\mathbf{x} = [x_1, x_2, x_3]$ and $\mathbf{y} = [y_1, y_2, y_3]$ be two vectors of R^3.
To check the linearity of a given function, we have to verify the following two properties:

(i) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$; (ii) $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$

$$\begin{aligned}
\text{(i)} \quad T(\mathbf{x} + \mathbf{y}) &= T(x_1 + y_1, x_2 + y_2, x_3 + y_3) \\
&= ((x_1 + y_1) + 2(x_3 + y_3), (x_1 + y_1) - (x_2 + y_2)) \\
&= (x_1 + 2x_3, x_1 - x_2) + (y_1 + 2y_3, y_1 - y_2) \\
&= T(\mathbf{x}) + T(\mathbf{y})
\end{aligned}$$

$$\begin{aligned}
\text{(ii)} \quad T(\alpha\mathbf{x}) &= T(\alpha x_1, \alpha x_2, \alpha x_3) \\
&= (\alpha x_1 + 2(\alpha x_3), \alpha x_1 - \alpha x_3) \\
&= \alpha(x_1 + 2x_3, x_1 - x_3) \\
&= \alpha T(\mathbf{x})
\end{aligned}$$

Hence the given function T is a linear transformation.

EXAMPLE 3.2 Check the linearity of the following transformation

$$T: R^3 \to R^3$$

defined by $T(x_1, x_2, x_3) = (x_1, x_2^2, x_3)$.

Solution: Let $\mathbf{x} = [x_1, x_2, x_3]$ and $\mathbf{y} = [y_1, y_2, y_3]$ be two vectors of R^3.

$$\begin{aligned}
\text{(i)} \qquad T(\mathbf{x} + \mathbf{y}) &= T(x_1 + y_1, x_2 + y_2, x_3 + y_3) \\
&= (x_1 + y_1, (x_2 + y_2)^2, x_3 + y_3)
\end{aligned}$$

(ii) $T(\mathbf{x}) + T(\mathbf{y}) = (x_1, x_2^2, x_3) + (y_1, y_2^2, y_3)$

$= (x_1 + y_1, x_2^2 + y_2^2, x_3 + y_3)$

Therefore, $T(\mathbf{x} + \mathbf{y}) \neq T(\mathbf{x}) + T(\mathbf{y})$. Hence the given function T is not a linear transformation.

EXAMPLE 3.3 Check the linearity of the following transformation

$$T: R^3 \to R^3$$

defined by $T(x_1, x_2, x_3) = (x_1 + 1, x_2 + 2, x_3 + 3)$.

Solution: Let $\mathbf{x} = [x_1, x_2, x_3]$, and $\mathbf{y} = [y_1, y_2, y_3]$ be two vectors of R^3.

(i) $T(\mathbf{x} + \mathbf{y}) = T(x_1 + y_1, x_2 + y_2, x_3 + y_3)$

$= (x_1 + y_1 + 1, x_2 + y_2 + 2, x_3 + y_3 + 3)$

(ii) $T(\mathbf{x}) + T(\mathbf{y}) = (x_1 + 1, x_2 + 2, x_3 + 3) + (y_1 + 1, y_2 + 2, y_3 + 3)$

$= (x_1 + y_1 + 2, x_2 + y_2 + 4, x_3 + y_3 + 6)$

$\neq T(\mathbf{x} + \mathbf{y})$

Therefore $T(\mathbf{x} + \mathbf{y}) \neq T(\mathbf{x}) + T(\mathbf{y})$. Hence the given transformation T is not a linear transformation.

EXAMPLE 3.4 Check the linearity of the following transformation

$$T: P_2 \to P_3$$

defined by $T(p(\mathbf{x})) = \mathbf{x}p(\mathbf{x})$.

Solution: Let $p(\mathbf{x})$, and $q(\mathbf{x})$ be the vectors of P_2.

(i) $T(p(\mathbf{x}) + q(\mathbf{x})) = \mathbf{x}(p(\mathbf{x}) + q(\mathbf{x}))$

$= \mathbf{x}p(\mathbf{x}) + \mathbf{x}q(\mathbf{x})$

$= T(p(\mathbf{x})) + T(q(\mathbf{x}))$

Thus $T(p(\mathbf{x}) + q(\mathbf{x})) = T(p(\mathbf{x})) + T(q(\mathbf{x}))$

(ii) $T(\alpha p(\mathbf{x})) = \mathbf{x}(\alpha p(\mathbf{x}))$

$= \alpha(\mathbf{x}p(\mathbf{x}))$

$= \alpha T(p(\mathbf{x}))$

Thus $T(\alpha p(\mathbf{x})) = \alpha T(p(\mathbf{x}))$.

Therefore the given $T: P_2 \to P_3$ is a linear transformation.

EXAMPLE 3.5 Check the linearity of the following transformation

$$T: P_2 \to P_2$$

defined by $T(p(\mathbf{x})) = p(\mathbf{x}) + c(\mathbf{x})$
where $c(x) = c_0 + c_1x + c_2x^2$ is a fixed vector in P_2.

Solution: Let $p(\mathbf{x})$ and $q(\mathbf{x})$ be the vectors of P_2.

(i) $T(p(\mathbf{x}) + q(\mathbf{x})) = (p(x) + q(x)) + c(x)$

$= p(x) + q(x) + c(x)$

(ii) $$T(p(\mathbf{x})) + T(q(\mathbf{x})) = (p(\mathbf{x}) + c(\mathbf{x})) + (q(\mathbf{x}) + c(\mathbf{x}))$$
$$= p(\mathbf{x}) + q(\mathbf{x}) + 2c(\mathbf{x})$$
$$\neq T(p(\mathbf{x}) + q(\mathbf{x}))$$

Therefore the given transformation T is not a linear transformation.

EXAMPLE 3.6 Check the linearity of the following transformation

$$T: M_{m\times m} \to M_{m\times n}$$

defined by $T(\mathbf{A}) = \mathbf{AB}$ where $\mathbf{B}$ is a fixed matrix of order $m \times n$.

Solution: Let $\mathbf{A}_1$ and $\mathbf{A}_2$ be two square matrices of order m
Since $\mathbf{B}$ is a fixed $m \times n$ matrix, therefore $\mathbf{A}_1\mathbf{B}$ and $\mathbf{A}_2\mathbf{B}$ are also matrices of order $m \times n$.

(i) $$T(\mathbf{A}_1 + \mathbf{A}_2) = (\mathbf{A}_1 + \mathbf{A}_2)\mathbf{B}$$
$$= \mathbf{A}_1\mathbf{B} + \mathbf{A}_2\mathbf{B}$$
$$= T(\mathbf{A}_1) + T(\mathbf{A}_2)$$

Thus $$T(\mathbf{A}_1 + \mathbf{A}_2) = T(\mathbf{A}_1) + T(\mathbf{A}_2).$$

(ii) $$T(\alpha\mathbf{A}_1) = (\alpha\mathbf{A}_1)\mathbf{B}$$
$$= \alpha(\mathbf{A}_1\mathbf{B}) = \alpha T(\mathbf{A}_1)$$

Thus $$T(\alpha\acute{\mathbf{A}}_1) = \alpha T(\mathbf{A}_1)$$

Hence the given transformation is a linear transformation from $M_{m\times m}$ to $M_{m\times n}$.

EXAMPLE 3.7 Check the linearity of the following transformation

$$T: M_{22} \to M_{22}$$

defined by $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a^2 + 2bc + d^2$

Solution: Let $\mathbf{A}_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$ and $\mathbf{A}_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$ be two vectors of M_{22}.

Since $$\mathbf{A}_1 + \mathbf{A}_2 = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{bmatrix} \in M_{22}$$

$$T(\mathbf{A}_1 + \mathbf{A}_2) = T\left(\begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{bmatrix}\right)$$
$$= (a_1 + a_2)^2 + 2(b_1 + b_2)\,(c_1 + c_2) + (d_1 + d_2)^2$$

Now $$T(\mathbf{A}_1) + T(\mathbf{A}_2) = T\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right)$$
$$= a_1^2 + 2b_1c_1 + d_1^2 + a_2^2 + 2b_2c_2 + d_2^2$$
$$= (a_1^2 + a_2^2) + 2(b_1c_1 + b_2c_2) + (d_1^2 + d_2^2)$$
$$\neq T(\mathbf{A}_1 + \mathbf{A}_2)$$

$$\therefore \qquad T(\mathbf{A}_1) + T(\mathbf{A}_2) \neq T(\mathbf{A}_1 + \mathbf{A}_2).$$

Therefore the given mapping $T: M_{22} \to M_{22}$ is not a linear transformation .

EXAMPLE 3.8 If $F(-\infty, \infty)$ is the vector space of real-valued functions defined on $(-\infty, \infty)$ and $C'(-\infty, \infty)$ is the vector space of real-valued functions with continuous first derivatives on $(-\infty, \infty)$, show that the function $T: C'(-\infty, \infty) \to F(-\infty, \infty)$ defined by

$$T(f(\mathbf{x}) = f'(\mathbf{x})$$

is a linear transformation function.

Solution: Let $f(\mathbf{x})$ and $g(\mathbf{x})$ be the vectors of $C'(-\infty, \infty)$.

(i)
$$\begin{aligned} T(f(\mathbf{x}) + g(\mathbf{x})) &= (f(\mathbf{x}) + g(\mathbf{x}))' \\ &= f'(\mathbf{x}) + g'(\mathbf{x}) \\ &= T(f(\mathbf{x})) + T(g(\mathbf{x})) \end{aligned}$$

Thus
$$T(f(\mathbf{x}) + g(\mathbf{x})) = T(f(\mathbf{x})) + T(g(\mathbf{x}))$$

(ii)
$$\begin{aligned} T(\alpha f(\mathbf{x})) &= (\alpha f(\mathbf{x}))' \\ &= \alpha f'(\mathbf{x}) \\ &= \alpha T(f(\mathbf{x})) \end{aligned}$$

Thus
$$T(\alpha f(\mathbf{x})) = \alpha T(f(\mathbf{x}))$$

Hence the given function is a linear transformation.

EXAMPLE 3.9 Let $C(-\infty, \infty)$ be a vector space of continuous functions defined on $(-\infty, \infty)$. Then the transformation $T: C(-\infty, \infty) \to C'(-\infty, \infty)$ is defined as

$$T(f(\mathbf{x})) = \int_0^x f(t)dt$$

Solution: Let $f(\mathbf{x})$ and $g(\mathbf{x})$ be two vectors of $(-\infty, \infty)$.

(i)
$$\begin{aligned} T((f+g)(\mathbf{x})) &= \int_0^x (f+g)(t)dt \\ &= \int_0^x ((f(t) + g(t))dt \\ &= \int_0^x f(t)dt + \int_0^x g(t)dt \\ &= T(f(\mathbf{x})) + T(g(\mathbf{x})) \end{aligned}$$

Thus
$$T(f(\mathbf{x}) + g(\mathbf{x})) = T(f(\mathbf{x})) + T(g(\mathbf{x})).$$

(ii)
$$T(\alpha f(\mathbf{x})) = \int_0^x \alpha f(t)dt$$

$$= \alpha \int_0^x f(t)dt$$

$$= \alpha T(f(\mathbf{x}))$$

Hence the given transformation is a linear transformation from $C(-\infty, \infty)$ to $C'(-\infty, \infty)$.

EXAMPLE 3.10 The transformation T: $C(-\infty, \infty) \to C(-\infty, \infty)$ is defined by

$$T(f(\mathbf{x})) = f(\mathbf{x}) + a$$

where a is a fixed nonzero scalar.
Is the transformation a linear transformation?

Solution: Let $f(\mathbf{x})$, $g(\mathbf{x})$ be two vectors of $C(-\infty, \infty)$, then

$$\begin{aligned} T(f(\mathbf{x}) + g(\mathbf{x})) &= f(\mathbf{x}) + g(\mathbf{x}) + a \\ T(f(\mathbf{x})) + T(g(\mathbf{x})) &= (f(\mathbf{x}) + a) + (g(\mathbf{x}) + a) \\ &= f(\mathbf{x}) + g(\mathbf{x}) + 2a \\ &\neq T(f(\mathbf{x}) + g(\mathbf{x})) \qquad \text{since } a \neq 0 \end{aligned}$$

Thus $$T(f(\mathbf{x}) + g(\mathbf{x})) \neq T(f(\mathbf{x})) + T(g(\mathbf{x}))$$

Therefore T is not a linear transformation from $C(-\infty, \infty)$ to $C(-\infty, \infty)$.

EXAMPLE 3.11 Let V and W be two vector spaces with the same set of scalars. Then show that the function T: $V \to W$ defined by $T(\mathbf{v}) = \mathbf{0}$ (zero transformation) for all $\mathbf{v}$ in V where $\mathbf{0}$ is a zero element of W, is a linear transformation.

Solution: Let $\mathbf{u}$ and $\mathbf{v}$ be the vectors of V. Then

$$T(\mathbf{u}) = \mathbf{0}, \qquad T(\mathbf{v}) = \mathbf{0}$$

(i) $T(\mathbf{u} + \mathbf{v}) = \mathbf{0}$ since $\mathbf{u} + \mathbf{v}$ is a vector of V

$$= \mathbf{0} + \mathbf{0}$$

$$= T(\mathbf{u}) + T(\mathbf{v})$$

(ii) $T(\alpha\mathbf{u}) = \mathbf{0}$ since $\alpha\mathbf{u}$ is a vector of V

$$= \alpha(\mathbf{0})$$

$$= \alpha T(\mathbf{u})$$

Hence the zero transformation is a linear transformation.

EXAMPLE 3.12 Let V be a vector space. If T: $V \to V$ is defined by $T(\mathbf{v}) = \mathbf{v}$ (identity transformation) for all $\mathbf{v}$ in V, then T is a linear transformation.

Solution: Let $\mathbf{u}$ and $\mathbf{v}$ be the vectors of V. Then

$$T(\mathbf{u}) = \mathbf{u}, \qquad T(\mathbf{v}) = \mathbf{v}$$

(i) $T(\mathbf{u} + \mathbf{v}) = \mathbf{u} + \mathbf{v}$ since $\mathbf{u} + \mathbf{v}$ is a vector of V

$$= T(\mathbf{u}) + T(\mathbf{v})$$

(ii) $T(\alpha\mathbf{u}) = \alpha\mathbf{u}$ since $\alpha\mathbf{u}$ is a vector of V

$$= \alpha T(\mathbf{u})$$

Hence the identity transformation is a linear transformation.

EXAMPLE 3.13 Consider the operator T: $R^2 \to R^2$ that maps each vector into its **orthogonal projection** on the x-axis (Figure 3.1), that is, $T(x, y) = (x, 0)$. Show that T is a linear operator on R^2.

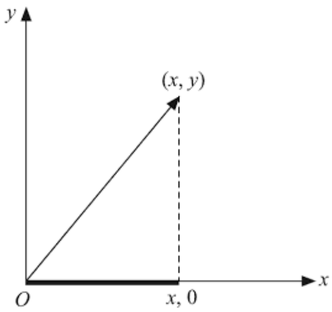


Figure 3.1 Orthogonal projection on x axis.

Solution: Let $\mathbf{u} = [x_1, y_1]$ and $\mathbf{v} = [x_2, y_2]$ be vectors of R^2 and α be any real scalar. Then

$$T(\mathbf{u}) = T(x_1, y_1) = (x_1, 0), \qquad T(\mathbf{v}) = T(x_2, y_2) = (x_2, 0)$$

(i)
$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T((x_1, y_1) + (x_2, y_2)) \\ &= T(x_1 + x_2, y_1 + y_2) \\ &= (x_1 + x_2, 0) \\ &= (x_1, 0) + (x_2, 0) \\ &= T(\mathbf{u}) + \mathrm{T}(\mathbf{v}) \end{aligned}$$

(ii)
$$\begin{aligned} T(\alpha\mathbf{u}) &= T(\alpha(x_1, y_1)) \\ &= T(\alpha x_1, \alpha y_1) \\ &= (\alpha x_1, 0) \\ &= \alpha(x_1, 0) \\ &= \alpha T(\mathbf{u}) \end{aligned}$$

Hence the orthogonal projection operator on x-axis in R^2 is a linear transformation on R^2.

EXAMPLE 3.14 Consider the operator T: $R^2 \to R^2$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y) = (0, y)$. Show that T is a linear operator. The verification of linearity of the operator is left to reader. See Example 3.13.

EXAMPLE 3.15 Consider the operator T: $R^3 \to R^3$ that maps each vector into its orthogonal projection on the x-axis, that is, $T(x, y, z) = (x, 0, 0)$. Show that T is a linear operator on R^3.

Solution: Let $\mathbf{u} = [x_1, y_1, z_1]$ and $\mathbf{v} = [x_2, y_2, z_2]$ be vectors of $R^3$and $\alpha \in R$. Then

$$T(\mathbf{u}) = T(x_1, y_1, z_1) = (x_1, 0, 0), \qquad T(\mathbf{v}) = T(x_2, y_2, z_2) = (x_2, 0, 0)$$

(i)
$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T((x_1, y_1, z_1) + (x_2, y_2, z_2)) \\ &= T(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_1 + x_2, 0, 0) \\ &= (x_1, 0, 0) + (x_2, 0, 0) \end{aligned}$$

$$= T(\mathbf{u}) + T(\mathbf{v})$$

(ii) $T(\alpha\mathbf{u}) = T(\alpha(x_1, y_1, z_1))$

$$= T(\alpha x_1, \alpha y_1, \alpha z_1)$$
$$= (\alpha x_1, 0, 0)$$
$$= \alpha(x_1, 0, 0)$$
$$= \alpha T(\mathbf{u})$$

Hence the orthogonal projection operator on x-axis in R^3 is a linear transformation on R^3.

EXAMPLE 3.16 Consider the operator $T: R^3 \to R^3$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y, z) = (0, y, 0)$. The reader can easily verify the linearity of T. See Example 3.15.

EXAMPLE 3.17 Consider the operator $T: R^3 \to R^3$ that maps each vector into its orthogonal projection on the z-axis, that is, $T(x, y, z) = (0, 0, z)$. The reader can easily verify that it is a linear operator on R^3. See Example 3.15.

EXAMPLE 3.18 Let V be an inner product space and W be a finite dimensional subspace of V (we will see in Chapter 4). Let $B = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots, \mathbf{w}_n\}$ be the orthonormal basis for the inner product space W. Consider the transformation $T: V \longrightarrow W$ that maps each vector into its orthogonal projection on the subspace W, that is,

$$T(\mathbf{u}) = \text{proj}_W \mathbf{u}$$
$$= \langle \mathbf{u}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{u}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u}, \mathbf{w}_n \rangle \mathbf{w}_n$$

shows that it is a linear transformation from V to W.

Solution: Let $\mathbf{u}, \mathbf{v}$ be two vectors of vector space V and let α be any real scalar. Then

$$T(\mathbf{u}) = \text{proj}_W \mathbf{u}$$
$$= \langle \mathbf{u}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{u}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u}, \mathbf{w}_n \rangle \mathbf{w}_n$$
$$T(\mathbf{v}) = \text{proj}_W \mathbf{v}$$
$$= \langle \mathbf{v}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{v}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{v}, \mathbf{w}_n \rangle \mathbf{w}_n$$

(i) $T(\mathbf{u} + \mathbf{v}) = \text{proj}_W (\mathbf{u} + \mathbf{v})$

$$= \langle \mathbf{u} + \mathbf{v}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{u} + \mathbf{v}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u} + \mathbf{v}, \mathbf{w}_n \rangle \mathbf{w}_n$$
$$= \left\{\left(\langle \mathbf{u}, \mathbf{w}_1 \rangle + \langle \mathbf{v}, \mathbf{w}_1 \rangle\right)\mathbf{w}_1\right\} + \left\{\left(\langle \mathbf{u}, \mathbf{w}_2 \rangle + \langle \mathbf{v}, \mathbf{w}_2 \rangle\right)\mathbf{w}_2\right\}$$
$$+ \cdots + \left\{\left(\langle \mathbf{u}, \mathbf{w}_n \rangle + \langle \mathbf{v}, \mathbf{w}_n \rangle\right)\mathbf{w}_n\right\}$$
$$= \left\{\langle \mathbf{u}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{u}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u}, \mathbf{w}_n \rangle \mathbf{w}_n\right\}$$
$$+ \left\{\langle \mathbf{v}, \mathbf{w}_1 \rangle \mathbf{w}_1 + \langle \mathbf{v}, \mathbf{w}_2 \rangle \mathbf{w}_2 + \cdots + \langle \mathbf{v}, \mathbf{w}_n \rangle \mathbf{w}_n\right\}$$
$$= T(\mathbf{u}) + T(\mathbf{v})$$

Thus $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii)
$$
\begin{aligned}
T(\alpha\mathbf{u}) &= \text{proj}_W(\alpha\mathbf{u}) \\
&= \langle \alpha\mathbf{u}, \mathbf{w}_1\rangle \mathbf{w}_1 + \langle \alpha\mathbf{u}, \mathbf{w}_2\rangle \mathbf{w}_2 + \cdots + \langle \alpha\mathbf{u}, \mathbf{w}_n\rangle \mathbf{w}_n \\
&= \alpha\langle \mathbf{u}, \mathbf{w}_1\rangle \mathbf{w}_1 + \alpha\langle \mathbf{u}, \mathbf{w}_2\rangle \mathbf{w}_2 + \cdots + \alpha\langle \mathbf{u}, \mathbf{w}_n\rangle \mathbf{w}_n \\
&= \alpha\left(\langle \mathbf{u}, \mathbf{w}_1\rangle \mathbf{w}_1 + \langle \mathbf{u}, \mathbf{w}_2\rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u}, \mathbf{w}_n\rangle \mathbf{w}_n\right) \\
&= \alpha T(\mathbf{u})
\end{aligned}
$$

Thus $T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$.

Hence, the orthogonal projection transformation from V to W is a linear transformation.

EXAMPLE 3.19 Let V be any vector space. Let T be the operator on V defined by

(i) $T(\mathbf{u}) = k\mathbf{u}$ where $k > 1$ is called the **Dilation** operator.
(ii) $T(\mathbf{u}) = k\mathbf{u}$ where $0 < k < 1$ is called the **Contraction** operator.

as shown in Figure 3.2.

Show that T is a linear operator on V.

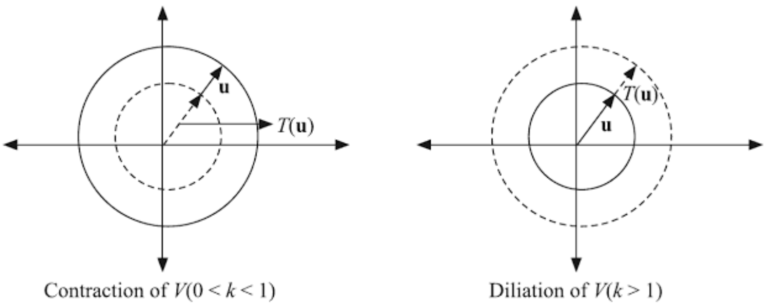


Figure 3.2

Solution: Let $\mathbf{u}, \mathbf{v}$ be two vectors of V and let α be any real scalar. Then

$$T(\mathbf{u}) = k\mathbf{u}, \qquad T(\mathbf{v}) = k\mathbf{v}$$

(i)
$$
\begin{aligned}
T(\mathbf{u}+\mathbf{v}) &= k(\mathbf{u}+\mathbf{v}) \\
&= k\mathbf{u} + k\mathbf{v} \\
&= T(\mathbf{u}) + T(\mathbf{v})
\end{aligned}
$$

Thus $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii)
$$
\begin{aligned}
T(\alpha\mathbf{u}) &= k(\alpha\mathbf{u}) \\
&= \alpha(k\mathbf{u}) \\
&= \alpha T(\mathbf{u})
\end{aligned}
$$

Thus $T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$

Hence the contraction and dilation operators are linear operators on V.

EXAMPLE 3.20

1. The transformation T: $R^2 \to R^2$ that maps each vector into its symmetric image about the x-axis (Figure 3.3) is called the **reflection operator** on R^2, that is,

$$T(x, y) = (x, -y)$$

Show that it is a linear operator on R^2.

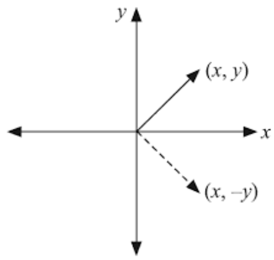


Figure 3.3 Reflection about x-axis.

Solution: Let $\mathbf{u} = [x_1, y_1]$ and $\mathbf{v} = [x_2, y_2]$ be two vectors of R^2 and $\alpha \in R$. Then

$$T(\mathbf{u}) = T(x_1, y_1) = (x_1, -y_1)$$
$$T(\mathbf{v}) = T(x_2, y_2) = (x_2, -y_2)$$

(i)
$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T(x_1 + x_2, y_1 + y_2) \\ &= (x_1 + x_2, -y_1 - y_2) \\ &= (x_1, -y_1) + (x_2, -y_2) \\ &= T(\mathbf{u}) + T(\mathbf{v}) \end{aligned}$$

Thus $\quad T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii)
$$\begin{aligned} T(\alpha\mathbf{u}) &= T(\alpha x_1, \alpha y_1) \\ &= (\alpha x_1, -\alpha y_1) = \alpha(x_1, -y_1) \\ &= \alpha T(\mathbf{u}) \end{aligned}$$

Thus $\quad T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$

Hence the reflection operator T: $R^2 \to R^2$ that maps each vector into its symmetric image about the x-axis is a linear transformation.

2. Similarly the reflection operator T: $R^2 \to R^2$ that maps each vector into its symmetric image about the y-axis (Figure 3.4) that is, $T(x, y) = (-x, y)$ is a linear transformation.

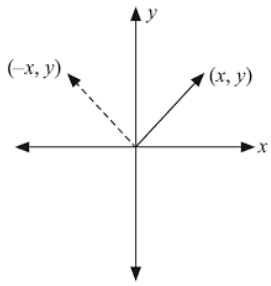


Figure 3.4 Reflection about y-axis.

EXAMPLE 3.21 The transformation T: $R^2 \to R^2$ that maps each vector into its symmetric image about the line $y = x$ (Figure 3.5) is a **reflection** operator (about the line $y = x$) in R^2, that is,

$$T(x, y) = (y, x)$$

The reader should verify linearity of T. (See Example 3.15)

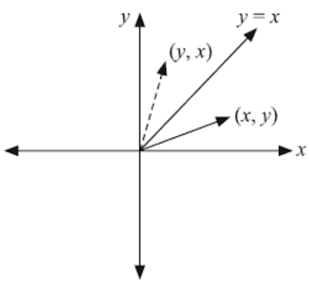


Figure 3.5 Reflection about $y = x$ line.

EXAMPLE 3.22 **1.** The transformation $T: R^3 \to R^3$ that maps each vector into its symmetric image about the xy plane (Figure 3.6) is a **reflection** operator (about the xy plane) on R^3, that is,

$$T(x, y, z) = (x, y, -z)$$

Show that it is a linear operator on R^3.

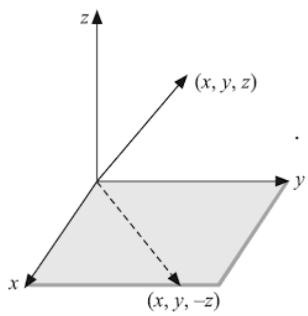


Figure 3.6 Reflection about xy plane.

Solution: Let $\mathbf{u} = [x_1, y_1, z_1]$ and $\mathbf{v} = [x_2, y_2, z_2]$ be two vectors in R^3.

$$T(\mathbf{u}) = T(x_1, y_1, z_1) = (x_1, y_1, -z_1)$$

$$T(\mathbf{v}) = T(x_2, y_2, z_2) = (x_2, y_2, -z_2)$$

(i)
$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_1 + x_2, y_1 + y_2, -z_1 - z_2) \\ &= (x_1, y_1, -z_1) + (x_2, y_2, -z_2) \\ &= T(\mathbf{u}) + T(\mathbf{v}) \end{aligned}$$

Thus $\quad T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii)
$$\begin{aligned} T(\alpha\mathbf{u}) &= T(\alpha x_1, \alpha y_1, \alpha z_1) \\ &= (\alpha x_1, \alpha y_1, -\alpha z_1) \\ &= \alpha(x_1, y_1, -z_1) \\ &= \alpha T(\mathbf{u}) \end{aligned}$$

Thus $\quad T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$

Hence the reflection operator (about the xy-plane) on R^3 is a linear operator on R^3.

2. The reflection operator about the xz-plane (Figure 3.7) on R^3 is defined by

$$T(x, y, z) = (x, -y, z)$$

The reader should verify the linearity of T. (See Example 3.22.)

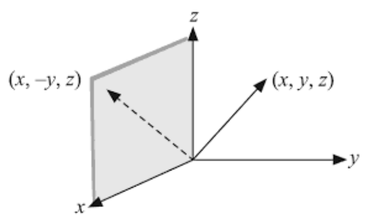


Figure 3.7 Reflection about xz plane.

3. The reflection operator about the yz-plane (Figure 3.8) on R^3 is defined by

$$T(x, y, z) = (-x, y, z)$$

The reader should verify the linearity of T. (See Example 3.22.)

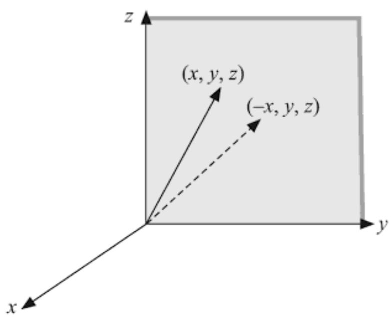


Figure 3.8 Reflection about yz plane.

EXAMPLE 3.23 The operator that rotates each vector through a fixed angle θ is called the **rotation operator** (Figure 3.9).

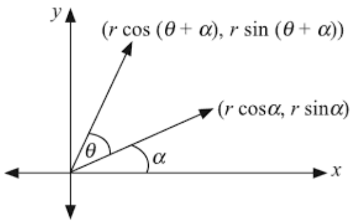


Figure 3.9 Rotation of the vector by an angle θ.

Derive the formula for a rotation operator in R^2 and show that it is a linear operator in R^2.

Solution: Let $\mathbf{u} = [x, y]$ be the vectors of R^2.

Here we will use the polar co-ordinates to derive the formulas. The polar co-ordinates of $\mathbf{u}$ are

$$x = r\cos\alpha, \qquad y = r\sin\alpha$$

If $\mathbf{u}$ rotates by a fixed angle θ, then the new co-ordinates (x', y') are given by the formulas,

$$\begin{aligned} x' &= r\cos(\alpha+\theta) & y' &= r\sin(\alpha+\theta)\\ &= r(\cos\alpha\cos\theta - \sin\alpha\sin\theta) & &= r(\sin\alpha\cos\theta + \sin\theta\cos\alpha)\\ &= (r\cos\alpha)\cos\theta - (r\sin\alpha)\sin\theta & &= (r\sin\alpha)\cos\theta + (r\cos\alpha)\sin\theta\\ &= x\cos\theta - y\sin\theta & &= y\cos\theta + x\sin\theta \end{aligned}$$

Therefore the rotation operator (rotates by an angle θ) on R^2 can be defined as

$$\begin{aligned} T(x, y) &= (x', y')\\ &= (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \end{aligned}$$

Here we want to verify linearity of the rotation operator on R^2.
Let $\mathbf{u} = [x_1, y_1]$ and $\mathbf{v} = [x_2, y_2] \in R^2$ and $k \in R$.

Then
$$\begin{aligned} T(\mathbf{u}) &= T(x_1, y_1)\\ &= (x_1\cos\theta - y_1\sin\theta, x_1\sin\theta + y_1\cos\theta)\\ T(\mathbf{v}) &= T(x_2, y_2)\\ &= (x_2\cos\theta - y_2\sin\theta, x_2\sin\theta + y_2\cos\theta) \end{aligned}$$

(i)
$$\begin{aligned} T(\mathbf{u}+\mathbf{v}) &= T(x_1+x_2, y_1+y_2)\\ &= ((x_1+x_2)\cos\theta - (y_1+y_2)\sin\theta, (x_1+x_2)\sin\theta + (y_1+y_2)\cos\theta)\\ &= ((x_1\cos\theta - y_1\sin\theta) + (x_2\cos\theta - y_2\sin\theta), (x_1\sin\theta + y_1\cos\theta)\\ &\quad + (x_2\sin\theta + y_2\cos\theta))\\ &= (x_1\cos\theta - y_1\sin\theta, x_1\sin\theta + y_1\cos\theta) + (x_2\cos\theta - y_2\sin\theta, x_2\sin\theta + y_2\cos\theta)\\ &= T(\mathbf{u}) + T(\mathbf{v}) \end{aligned}$$

Thus $\quad T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$

(ii) $\quad T(k\mathbf{u}) = T(kx_1, ky_1)$

$$= ((kx_1)\cos\theta - (ky_1)\sin\theta, (kx_1)\sin\theta + (ky_1)\cos\theta)$$

$$= k(x_1\cos\theta - y_1\sin\theta, x_1\sin\theta + y_1\cos\theta)$$

$$= kT(\mathbf{u})$$

Thus $\quad T(k\mathbf{u}) = kT(\mathbf{u})$.

Hence, the rotation operator on R^2 is a linear operator on R^2.

EXAMPLE 3.24 Derive the formula for a rotation operator on R^3 that rotates each vector through a fixed angle θ about the positive z-axis in the counterclockwise direction (Figure 3.10). Verify the linearity of such operator.

Solution: Let $\mathbf{u} = [x, y, z]$ be any vector of R^3.

If $\mathbf{u}$ rotates about z-axis, then its z co-ordinates do not change and xy co-ordinates are changed by an angle θ.

Therefore the new co-ordinates (x', y', z') are

$$x' = x\cos\theta - y\sin\theta$$

$$y' = x\sin\theta + y\cos\theta$$

$$z' = z$$

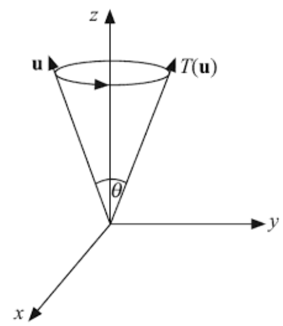


Figure 3.10 Rotation about z-axis through an angle θ.

Hence, the rotation operator (about z-axis) on R^3 is defined by

$$T(x, y, z) = (x', y', z')$$

$$= (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z)$$

It is easy to verify that T is a linear operator. (This is left to the reader. See Example 3.23.)

EXAMPLE 3.25 Derive the formula for a rotation operator on R^3 that rotates each vector through a fixed angle θ about the positive y-axis in the counterclockwise direction (Figure 3.11). Show that it is a linear operator on R^3.

Solution: Let $\mathbf{u} = [x, y, z]$ be any vector of R^3.

If $\mathbf{u}$ rotates about y-axis, then its y-co-ordinate does not change and zx-co-ordinates are changed by an angle θ. Therefore, the new co-ordinates (x', y', z') are

$$y' = y$$

$$z' = z\cos\theta - x\sin\theta$$

$$x' = z\sin\theta + x\cos\theta$$

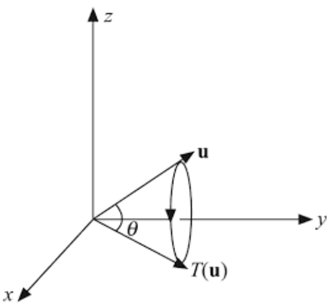


Figure 3.11 Rotation about y-axis through an angle θ.

i.e. $\quad T(x, y, z) = (x', y', z')$

$$= (z\cos\theta - x\sin\theta, y, z\sin\theta + x\cos\theta)$$

Hence T is a linear operator on R^3 can be verified as in Example 3.23.

EXAMPLE 3.26 Derive the formula for a rotation operator on R^3 that rotates each vector through a fixed angle θ about the positive x-axis in the anticlockwise direction (Figure 3.12). Show that it is a linear operator on R^3.

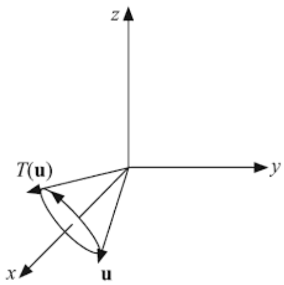


Figure 3.12 Rotation about the x-axis through an angle θ.

Solution: Let $\mathbf{u} = [x, y, z]$ be any vector of R^3.

If $\mathbf{u}$ rotates about the x-axis, then its x co-ordinate does not change and the yz co-ordinates are changed by an angle θ.

Therefore, the new co-ordinates (x', y', z') are

$$x' = x$$

$$y' = y\cos\theta - z\sin\theta$$

$$z' = z\cos\theta + y\sin\theta$$

i.e.
$$T(x, y, z) = (x', y', z')$$
$$= (x, y\cos\theta - z\sin\theta, z\cos\theta + y\sin\theta)$$

Hence, T is a linear operator on R^3 can be verified as in Example 3.22.

Theorem 3.1 [Linear Transformation]

Let $T: V \to W$ be a linear transformation. If $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n$ are vectors in V and $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ are scalars, then

$$T(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 + \cdots + \alpha_n\mathbf{x}_n) = \alpha_1 T(\mathbf{x}_1) + \alpha_2 T(\mathbf{x}_2) + \alpha_3 T(\mathbf{x}_3) + \cdots + \alpha_n T(\mathbf{x}_n).$$

Corollary 3.1 Let $T: V \to W$ be a linear transformation and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ be a basis for V.

If $\qquad \mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \cdots + \alpha_n\mathbf{v}_n$ is a vector of V

then $\qquad T(\mathbf{v}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \alpha_3 T(\mathbf{v}_3) + \cdots + \alpha_n T(\mathbf{v}_n). \in W$

Remark: The above corollary says that if we have the images $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ of the basis vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$, then we can find the image of any vector $\mathbf{v}$ in V under the linear transformation T.

Corollary 3.2 [Properties of Linear Transformation]

If $T: V \to W$ is a linear transformation, then

(i) $T(\mathbf{0}) = \mathbf{0}$ $\qquad$ (ii) $T(-\mathbf{u}) = -T(\mathbf{u})$

(iii) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ $\qquad$ for all $\mathbf{u}$ and $\mathbf{v}$ in V.

EXAMPLE 3.27 Determine the linear transformation T of the following whose images of basis vectors are given as follows:

(i) $T: R^2 \to R^2$, basis $B = \{[1, 0], [0, 1]\}$

$T(1, 0) = (1, 1)$ and $T(0, 1) = (2, -1)$. Also, compute $T(5, -2)$.

(ii) $T: R^3 \to R^3$, basis $B = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$
$T(1, 0, 0) = (1, -1, 1)$, $T(0, 1, 0) = (1, 0, 2)$, $T(0, 0, 1) = (2, 3, 0)$.
Also, compute $T(1, -3, 2)$.

(iii) $T: R^2 \to R^2$, basis $B = \{[2, -4], [3, 8]\}$

$T(2, -4) = (1, 2)$, $\quad T(3, 8) = \left(\frac{3}{2}, -4\right)$

Also, compute $T(4, 6)$.

(iv) $T: R^3 \to R^2$, basis $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$
$T(1, 0, 0) = (2, 0)$, $T(1, 1, 0) = (1, 1)$, $T(1, 1, 1) = (2, -3)$.
Also, compute $T(1, 5, 2)$.

Solution:

(i) Here basis $B = \{[1, 0], [0, 1]\}$ of R^2

$$T(1, 0) = (1, 1),\ T(0, 1) = (2, -1)$$

We want to find the formula of T for the general vector $[x_1, x_2]$ in R^2.
For this, first we represent $[x_1, x_2]$ as a linear combination of $[1, 0]$ and $[0, 1]$ and then we will use Corollary 3.1 to find the formula for T.

$$[x_1, x_2] = \alpha_1[1, 0] + \alpha_2[0, 1] \qquad \alpha_1, \alpha_2 \in R$$

$$[x_1, x_2] = [\alpha_1, \alpha_2]$$

$\therefore$ $\quad \alpha_1 = x_1,\ \alpha_2 = x_2$

and $\quad [x_1, x_2] = x_1[1, 0] + x_2[0, 1]$

By applying Corollary 3.1, we get

$$\begin{aligned} T(x_1, x_2) &= x_1 T(1, 0) + x_2 T(0, 1) \\ &= x_1(1, 1) + x_2(2, -1) \\ &= (x_1 + 2x_2, x_1 - x_2) \end{aligned}$$

Thus, $T(x_1, x_2) = (x_1 + 2x_2, x_1 - x_2)$. From this formula, $T(5, -2) = (1, 7)$

(ii) Here basis of R^3 is $B = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$. The image of basis vectors are given as $T(1, 0, 0) = (1, -1, 1)$, $T(0, 1, 0) = (1, 0, 2)$, $T(0, 0, 1) = (2, 3, 0)$. We want to find the formula $T(x_1, x_2, x_3)$.

For this, we have to first find the expression for $[x_1, x_2, x_3]$ as a linear combination of basis vectors.

$$[x_1, x_2, x_3] = \alpha_1[1, 0, 0] + \alpha_2[0, 1, 0] + \alpha_3[0, 0, 1]$$

$$[x_1, x_2, x_3] = [\alpha_1, \alpha_2, \alpha_3]$$

$\therefore$ $\quad \alpha_1 = x_1, \quad \alpha_2 = x_2, \quad \alpha_3 = x_3$

and $\quad [x_1, x_2, x_3] = x_1[1, 0, 0] + x_2[0, 1, 0] + x_3[0, 0, 1]$

By applying Corollary 3.1, we get

$$\begin{aligned} T(x_1, x_2, x_3) &= x_1 T(1, 0, 0) + x_2 T(0, 1, 0) + x_3 T(0, 0, 1) \\ &= x_1(1, -1, 1) + x_2(1, 0, 2) + x_3(2, 3, 0) \end{aligned}$$

$$= (x_1 + x_2 + 2x_3, -x_1 + 3x_3, x_1 + 2x_2)$$

From this formula, $T(1, -3, 2) = (2, 5, -5)$

(iii) The given basis for R^2 is $B = \{[2, -4], [3, 8]\}$ and its images under T are

$$T(2, -4) = (1, 2), \qquad T(3, 8) = \left(\frac{3}{2}, -4\right)$$

First we try to express (x_1, x_2) as a linear combination of basis vectors $[2, -4]$ and $[3, 8]$. For that, we consider

$$[x_1, x_2] = \alpha_1[2, -4] + \alpha_2[3, 8]$$
$$[x_1, x_2] = [2\alpha_1 + 3\alpha_2, -4\alpha_1 + 8\alpha_2]$$

$\therefore$ $$x_1 = 2\alpha_1 + 3\alpha_2, \qquad x_2 = -4\alpha_1 + 8\alpha_2$$

By using Corollary 3.1, we have

$$\begin{aligned}
T(x_1, x_2) &= \frac{1}{28}(8x_1 - 3x_2)\, T(2, -4) + \frac{1}{14}(2x_1 + x_2)\, T(3, 8) \\
&= \frac{1}{28}(8x_1 - 3x_2)\,(1, 2) + \frac{1}{14}(2x_1 + x_2)\left(\frac{3}{2}, -4\right) \\
&= \left(\frac{8x_1 - 3x_2}{28} + \frac{6x_1 + 3x_2}{28},\ \frac{8x_1 - 3x_2}{14} + \frac{-8x_1 - 4x_2}{14}\right) \\
&= \left(\frac{x_1}{2}, \frac{-x_2}{2}\right) \\
&= \left(\frac{x_1}{2}, \frac{-x_2}{2}\right)
\end{aligned}$$

Thus, required linear transformation is

$$T(x_1, x_2) = \left(\frac{x_1}{2}, \frac{-x_2}{2}\right)$$

Using the above formula, we can obtain $T(4, 6) = \left(\frac{4}{2}, \frac{-6}{2}\right) = (2, -3)$

(iv) The basis for R^3 is $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$.
The images of basis vectors under a linear transformation $T: R^3 \to R^2$ are

$$T(1, 0, 0) = (2, 0), \quad T(1, 1, 0) = (1, 1), \quad T(1, 1, 1) = (2, -3)$$

Let $[x_1, x_2, x_3]$ be any vector in R^3. Then

$$[x_1, x_2, x_3] = \alpha_1[1, 0, 0] + \alpha_2[1, 1, 0] + \alpha_3[1, 1, 1]$$
$$[x_1, x_2, x_3] = [\alpha_1 + \alpha_2 + \alpha_3, \alpha_2 + \alpha_3, \alpha_3]$$

$\therefore$ $$\alpha_3 = x_3, \ \alpha_2 = x_2 - x_3, \ \alpha_1 = x_1 - x_2$$

and $$[x_1, x_2, x_3] = (x_1 - x_2)\,[1, 0, 0] + (x_2 - x_3)\,[1, 1, 0] + x_3\,[1, 1, 1]$$

Now, we apply

$$T(x_1, x_2, x_3) = (x_1 - x_2)\, T(1, 0, 0) + (x_2 - x_3)\, T(1, 1, 0) + x_3\, T(1, 1, 1)$$
$$= (x_1 - x_2)\,(2, 0) + (x_2 - x_3)\,(1, 1) + x_3\,(2, -3)$$
$$= (2x_1 - x_2 + x_3, x_2 - 4x_3)$$

Thus $\quad T(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, x_2 - 4x_3)$

Hence $\quad T(1, 5, 2) = (-1, -3)$

EXAMPLE 3.28 Let $T : V \to W$ be a linear transformation and Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ be a basis for V.

(i) If $T(\mathbf{v}_1) = T(\mathbf{v}_2) = \cdots = T(\mathbf{v}_n) = 0$, then T is a zero transformation.
(ii) If $T(\mathbf{v}_n) = \mathbf{v}_n$, then T is an identity transformation.

Solution: The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ is a basis for V (given).
Let $\mathbf{u}$ be any vector in V, then $\mathbf{u}$ can be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$

(i) The map $T: V \to W$ is a linear transformation and

$$T(\mathbf{v}_1) = T(\mathbf{v}_2) = \cdots = T(\mathbf{v}_n) = \mathbf{0}$$

By Corollary 3.1,

$$T(\mathbf{u}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \cdots + \alpha_n T(\mathbf{v}_n)$$
$$= \alpha_1(\mathbf{0}) + \alpha_2(\mathbf{0}) + \cdots + \alpha_n(\mathbf{0})$$
$$= \mathbf{0}$$
$$T(\mathbf{u}) = \mathbf{0}, \text{ for all } \mathbf{u} \in V$$

Hence T is a zero transformation.

(ii) The map $T: V \to W$ is a linear transformation and

$$T(\mathbf{v}_i) = \mathbf{v}_i \qquad \text{for each } i,\ 1 \le i \le n$$

By Corollary 3.1,

$$T(\mathbf{u}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \cdots + \alpha_n T(\mathbf{v}_n)$$
$$= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$
$$= \mathbf{u}$$
$$T(\mathbf{u}) = \mathbf{u}, \qquad \text{for all } \mathbf{u} \in V$$

Hence T is an identity transformation.

EXERCISE SET 1

1. Check the linearity of the following functions:

(i) $T: M_n \to R$, $\quad T(\mathbf{A}) = \det \mathbf{A}$
(ii) $T: V \to R^n$, $\quad (\mathbf{u})_s = k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_n\mathbf{u}_n$, $[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ basis for V, then $T(\mathbf{u}) = (\mathbf{u})_s$, $(k_1, k_2, k_3, \ldots, k_n)\ \varepsilon R^n$
(iii) $T: P_n \to P_n$, $\quad T(p(\mathbf{x})) = p(a\mathbf{x} + b)$, $\ a, b$ are scalars.

2. Check the linearity of the following functions:

(i) The operator T: $R^2 \to R^2$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y) = (0, y)$ is a linear operator

(ii) The operator T: $R^3 \to R^3$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y, z) = (0, y, 0)$.

(iii) The operator T: $R^3 \to R^3$ that maps each vector into its orthogonal projection on the z-axis, that is, $T(x, y, z) = (0, 0, z)$. It is a linear operator on R^3.

3. Check the linearity of the following functions:

(i) The transformation T: $R^2 \to R^2$ that maps each vector into its symmetric image about the line $y = x$ is a *reflection* operator (about the line $y = x$) R^2, that is, $T(x, y) = (y, x)$.

(ii) The reflection operator about the xz-plane on R^3 is defined by $T(x, y, z) = (x, -y, z)$.

(iii) The reflection operator about the yz-plane on R^3 is defined by $T(x, y, z) = (-x, y, z)$.

4. Check the linearity of the following functions:

(i) Derive the formula for a rotation operator on R^3 that rotates each vector through a fixed angle θ about the positive y-axis in the counterclockwise direction. Show that it is a linear operator on R^3.

(ii) Derive the formula for a rotation operator on R^3 that rotates each vector through a fixed angle θ about the positive x-axis in the anticlockwise direction and show that it is a linear operator on R^3.

5. Determine the linear transformation T of the following whose images of basis vectors are given.

(i) T: $R^3 \to R^3$, $B = \{[1, 2, 1], [2, 9, 0], [3, 3, 4]\}$ is a basis for R^3, $T(1, 2, 1) = (2, 0, 3)$, $T(2, 9, 0) = (2, 0, 11)$, $T(3, 3, 4) = (7, 0, 6)$. Hence compute $T(1, 0, 1)$.

(ii) T: $R^2 \to R^3$, $B = \{[1, 1], [0, 2]\}$ is a basis for R^2, $T(1, 1) = (1, 1, 2)$, $T(0, 2) = (0, 2, 2)$. Hence compute $T(2, 3)$.

(iii) T: $P_2 \to R$, $B = \{1, x, x^2\}$ is a basis for P_2, $T(1) = 1$, $T(\mathbf{x}) = 2$, $T(\mathbf{x}^2) = 3$. Hence compute $T(5 - 3x + 2x^2)$.

(iv) T: $R^2 \to R^2$, $B = \{[1, 0], [0, 1]\}$ is a basis for R^2, $T(1, 0) = (2, 1)$, $T(0, 1) = (-1, 1)$. Hence compute $T(3, 3)$.

(v) T: $R^3 \to R^3$, $B = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ is a basis for R^3, $T(1, 0, 0) = (1, 0, 1)$, $T(0, 1, 0) = (1, 2, -1)$, $T(0, 0, 1) = (1, 0, -1)$. Hence compute $T(1, -2, 1)$.

(vi) $T: M_{22} \to R$, $B = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, $T\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\} = 3$, $T\left\{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\} = -4$,

$T\left\{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\} = 1$, $T\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\} = -1$.

Hence compute $T\left\{\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}\right\}$.

3.3 MATRIX ASSOCIATED WITH A LINEAR TRANSFORMATION

Let V and W be two vector spaces of dimensions n and m respectively. Suppose $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are bases for the vector spaces V and W respectively. We have seen in the previous

section that to find the linear transformation T from V to W, it is sufficient to find the image $T(\mathbf{e}_k)$ of each basis vector $\mathbf{e}_k$ $(1 \le k \le n)$. Each vector $T(\mathbf{e}_k)$ in W can be expressed uniquely as a linear combination of the basis vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$, say

$$T(\mathbf{e}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{w}_i; \qquad 1 \le k \le n$$

or in ordered pair

$$[T(\mathbf{e}_k)]_{B'} = [a_{1k}, a_{2k}, \ldots, a_{mk}]; \qquad 1 \le k \le n$$

$$= \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}_{B'}$$

Similarly, each of n elements $T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_k)$ has one column vector. If each of them is considered a column matrix, then it forms an $m \to n$ ordered matrix as follows:

$$\mathbf{A} = [[T(\mathbf{e}_1)]_{B'}\ [T(\mathbf{e}_2)]_{B'} \cdots\ [T(\mathbf{e}_n)]_{B'}]$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

This matrix $\mathbf{A}$ is called an associated matrix of the linear transformation T from V to W. It is denoted by

$$\mathbf{A} = [T]_{B',B}$$

Remark: Here B' is a basis for the co-domain vector space W of T and B is a basis for domain vector space V.

For the Euclidean Vector Spaces

Recall the definition of linear transformation T from R^n to R^m,

$$T(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$$

where

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{x}_1 + a_{12}\mathbf{x}_2 + \cdots + a_{1n}\mathbf{x}_n \\ \mathbf{u}_2 &= a_{21}\mathbf{x}_1 + a_{22}\mathbf{x}_2 + \cdots + a_{2n}\mathbf{x}_n \\ \vdots \quad & \quad \vdots \qquad\quad \vdots \qquad \cdots \qquad \vdots \\ \mathbf{u}_m &= a_{m1}\mathbf{x}_1 + a_{m2}\mathbf{x}_2 + \cdots + a_{mn}\mathbf{x}_n \end{aligned}$$

Suppose $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ are the standard basis for V and W respectively. Then

$$\begin{aligned} T(\mathbf{e}_k) &= T(0, 0, \ldots, 1, 0, \ldots, 0) \\ &= [a_{1k}, a_{2k}, \ldots, a_{mk}] \end{aligned}$$

$$[T(\mathbf{e}_k)]_{B'} = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}$$

Therefore the associated matrix of linear transformation $T: R^n \to R^m$ is

$$\mathbf{A} = [T]_{B',B}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Remark: $[T]_{B',B}$ of linear transformation $T: V \to W$ has the property

$$[T(\mathbf{x})]_{B'} = [T]_{B',B}\,[\mathbf{x}]_B$$

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}_{B'} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Algorithm to Find the Matrix Associated to Linear Transformation

Let V and W be vector spaces. Suppose the sets $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are bases for the vector spaces V and W respectively and T is linear transformation from V to W. Then the matrix associated with the linear transformation can be calculated using the following algorithm.

Step 1 Find the images of basis vectors of B under T, that is, find $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, …, $T(\mathbf{e}_n)$

Step 2 Find the co-ordinates of images with respect to basis B', that is, express the images as a linear combinations of vectors of basis B'. In other words,

$$T(\mathbf{e}_k) = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

$$[T(\mathbf{e}_k)]_{B'} = [\alpha_1, \alpha_2, \ldots, \alpha_n]$$

Find the αs for each k, $1 \le k \le n$.

Step 3 Put them as columns of an $m \times n$ matrix

$$\mathbf{A} = [[T(\mathbf{e}_1)]_{B'}\ [T(\mathbf{e}_2)]_{B'} \ldots [T(\mathbf{e}_n)]_{B'}]$$

EXAMPLE 3.29 Find the associated matrix of the following linear transformation.

$T: R^2 \to R^2$, $T(x_1, x_2) = (2x_1 + x_2, x_1 - x_2)$ with bases $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ for the domain and co-domain of T.

Solution: Here $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is a standard basis of R^2. To find the matrix of a given linear transformation, first we have to find the images of basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$ under T and then find the co-ordinates of the images with respect to basis B of co-domain space

$$T(\mathbf{e}_1) = T(1, 0) = (2, 1) \qquad T(\mathbf{e}_2) = T(0, 1) = (1, -1)$$

Since the basis vectors of co-domain space are $\mathbf{e}_1$ and $\mathbf{e}_2$,

$$T(\mathbf{e}_1) = 2(1, 0) + 1(0, 1) = 2\mathbf{e}_1 + 1\mathbf{e}_2 \qquad T(\mathbf{e}_2) = 1(1, 0) - 1(0, 1) = 1\mathbf{e}_1 + 2\mathbf{e}_2$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Therefore the associated matrix of a linear transformation T is

$$\mathbf{A} = [[T(\mathbf{e}_1)]_B\ [T(\mathbf{e}_2)]_B$$
$$= \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

EXAMPLE 3.30 Find the associated matrix of the following linear transformation.

T: $R^2 \to R^2$, $T(x_1, x_2) = (2x_1 + x_2, x_1 - x_2)$ with bases $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ and $B' = \{[1, 1], [0, 2]\}$ for the domain and co-domain of T respectively.

Solution: Here $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is a standard basis of R^2. To find the matrix of a given linear transformation, first we have to find the images of basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$ under T and then find the co-ordinates of the images with respect to basis B' of co-domain space

$$T(\mathbf{e}_1) = T(1, 0) = (2, 1) \qquad T(\mathbf{e}_2) = T(0, 1) = (1, -1)$$

Since [1, 1] and [2, 1] are the basis vectors of co-domain space of T. So we will try to find the co-ordinates of $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$ with respect to basis $B' = \{[1, 1], [0, 2]\}$. Consider the equation,

$$T(\mathbf{e}_1) = \alpha_1(1, 1) + \alpha_2(0, 2) \qquad T(\mathbf{e}_2) = \beta_1(1, 1) + \beta_2(0, 2)$$
$$(2, 1) = (\alpha_1, \alpha_1 + 2\alpha_2) \qquad (1, -1) = (\beta_1, \beta_1 + 2\beta_2)$$
$$\alpha_1 = 2, \quad \alpha_2 = -\frac{1}{2} \qquad \beta_1 = 1, \quad \beta_2 = -1$$
$$T(\mathbf{e}_1) = 2(1, 1) + -\frac{1}{2}(0, 2) \qquad T(\mathbf{e}_2) = 1(1, 1) + (-1)(0, 2)$$
$$[T(\mathbf{e}_1)]_{B'} = \begin{bmatrix} 2 \\ -1/2 \end{bmatrix} \qquad [T(\mathbf{e}_2)]_{B'} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Therefore the associated matrix of a linear transformation T is

$$\mathbf{A} = [[T(\mathbf{e}_1)]_{B'}\ [T(\mathbf{e}_2)]_{B'}]$$
$$= \begin{bmatrix} 2 & 1 \\ -1/2 & -1 \end{bmatrix}$$

EXAMPLE 3.31 Find the associated matrix of the following linear transformation.

T: $R^3 \to R^3$, $T(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 + x_3)$ with bases $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$ and $B' = \{[1, 0, 1], [0, 1, 0], [1, 0, -1]\}$ for the domain and co-domain of T respectively.

Solution: Here $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$ is standard basis of R^3. To find the matrix of a given linear transformation, first we have to find the images of vectors of basis B under T and then find the co-ordinates of the images with respect to basis B' of co-domain space.

$$T(1, 0, 0) = (1, 0, 1); \qquad T(1, 1, 0) = (2, 1, 1); \qquad T(1, 1, 1) = (2, 2, 2)$$

Since [1, 0, 1], [0, 1, 0] and [1, 0, –1] are the basis vectors of co-domain space of T, we will try to find the co-ordinates of $T(1, 0, 0)$, $T(1, 1, 0)$, $T(1, 1, 1)$ with respect to basis B'. Consider the equation,

$$T(1, 0, 0) = \alpha_1(1, 0, 1) + \alpha_2(0, 1, 0) + \alpha_3(1, 0, -1)$$
$$T(1, 1, 0) = \beta_1(1, 0, 1) + \beta_2(0, 1, 0) + \beta_3(1, 0, -1)$$
$$T(1, 1, 1) = \gamma_1(1, 0, 1) + \gamma_2(0, 1, 0) + \gamma_3(1, 0, -1)$$

$$(1, 0, 1) = (\alpha_1 + \alpha_3, \alpha_2, \alpha_1 - \alpha_3); \quad (2, 1, 1) = (\beta_1 + \beta_3, \beta_2, \beta_1 - \beta_3); \quad (2, 2, 2) = (\gamma_1 + \gamma_3, \gamma_2, \gamma_1 - \gamma_3)$$

$$\alpha_1 = 1, \quad \alpha_2 = 0, \quad \alpha_3 = 0; \qquad \beta_1 = \frac{3}{2}, \quad \beta_2 = 1, \quad \beta_3 = \frac{1}{2}; \qquad \gamma_1 = 2, \quad \gamma_2 = 2, \quad \gamma_3 = 0$$

$$[T(1, 0, 0)]_{B'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(1, 1, 0)]_{B'} = \begin{bmatrix} 3/2 \\ 1 \\ 1/2 \end{bmatrix}; \qquad [T(1, 1, 1)]_{B'} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$$

Therefore the associated matrix of a linear transformation T is

$$\mathbf{A} = [[T(1, 0, 0)]_{B'}\ [T(1, 1, 0)]_{B'}\ [T(1, 1, 1)]_{B'}]$$
$$= \begin{bmatrix} 1 & 3/2 & 2 \\ 0 & 1 & 2 \\ 0 & 1/2 & 0 \end{bmatrix}$$

EXAMPLE 3.32 Find the associated matrix of the following linear transformation.

$T: R^3 \to R^2$, $T(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, x_2 - 4x_3)$ with bases $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$ and $B' = \{[1, 1], [1, -1]\}$ for the domain and co-domain of T respectively.

Solution: Here $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$ is standard basis of R^3. To find the matrix of a given linear transformation, first we have to find the images of vectors of basis B under T and then find the co-ordinates of the images with respect to basis B' of co-domain space.

$$T(1, 0, 0) = (2, 0); \qquad T(1, 1, 0) = (1, 1); \qquad T(1, 1, 1) = (2, -3)$$

Since [1, 1] and [1, –1] are the basis vectors of co-domain space of T, we will try to find the co-ordinates of $T(1, 0, 0)$, $T(1, 1, 0)$, $T(1, 1, 1)$ with respect to basis B'. Consider the equation

$$T(1, 0, 0) = \alpha_1(1, 1) + \alpha_2(1, -1)$$
$$T(1, 1, 0) = \beta_1(1, 1) + \beta_2(1, -1)$$
$$T(1, 1, 1) = \gamma_1(1, 1) + \gamma_2(1, -1)$$

$$(2, 0) = (\alpha_1 + \alpha_2, \alpha_1 - \alpha_2); \qquad (1, 1) = (\beta_1 + \beta_2, \beta_1 - \beta_2); \qquad (2, -3) = (\gamma_1 + \gamma_2, \gamma_1 - \gamma_2)$$

$$\alpha_1 = 1, \quad \alpha_2 = 1; \qquad \beta_1 = 1, \quad \beta_2 = 0; \qquad \gamma_1 = -\frac{1}{2}, \ \gamma_2 = \frac{5}{2}$$

$$[T(1, 0, 0)]_{B'} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \qquad [T(1, 1, 0)]_{B'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad [T(1, 1, 1)]_{B'} = \begin{bmatrix} -1/2 \\ 5/2 \end{bmatrix};$$

Therefore the associated matrix of a linear transformation T is

$$\mathbf{A} = [[T(1, 0, 0)]_{B'}\ [T(1, 1, 0)]_{B'}\ [T(1, 1, 1)]_{B'}]$$
$$= \begin{bmatrix} 1 & 1 & -1/2 \\ 1 & 0 & 5/2 \end{bmatrix}$$

EXAMPLE 3.33 Let T: $R^2 \to R^2$ be a linear transformation such that

$$T(\mathbf{e}_1 + \mathbf{e}_2) = 3\mathbf{e}_1 + 9\mathbf{e}_2; \qquad T(3\mathbf{e}_1 + 2\mathbf{e}_2) = 7\mathbf{e}_1 + 23\mathbf{e}_2$$

where $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for R^2.

(i) Determine the matrix of T relative to the given basis B.

(ii) Use the basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ for domain space R^2, find a new basis of the form $\{\mathbf{e}_1 + a\mathbf{e}_2, 2\mathbf{e}_1 + b\mathbf{e}_2\}$ for co-domain space R^2 relative to which the matrix of T will be in diagonal form.

Solution: Here $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for domain space R^2 and T is a linear transformation with

$$T(\mathbf{e}_1 + \mathbf{e}_2) = 3\mathbf{e}_1 + 9\mathbf{e}_2; \qquad T(3\mathbf{e}_1 + 2\mathbf{e}_2) = 7\mathbf{e}_1 + 23\mathbf{e}_2$$

Since T is a linear transformation,

$$T(\mathbf{e}_1) + T(\mathbf{e}_2) = 3\mathbf{e}_1 + 9\mathbf{e}_2; \qquad 3T(\mathbf{e}_1) + 2T(\mathbf{e}_2) = 7\mathbf{e}_1 + 23\mathbf{e}_2$$
$$T(\mathbf{e}_1) = \mathbf{e}_1 + 5\mathbf{e}_2; \qquad T(\mathbf{e}_2) = 2\mathbf{e}_1 + 4\mathbf{e}_2$$
$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 1 \\ 5 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

(i) The matrix of T relative to the given basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is

$$\mathbf{A} = [T]_{B,B}$$
$$= [[T(\mathbf{e}_1)]_B\ [T(\mathbf{e}_2)]_B]$$
$$= \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$$

(ii) Suppose $B' = \{\mathbf{e}_1 + a\mathbf{e}_2, 2\mathbf{e}_1 + b\mathbf{e}_2\}$ is a basis for co-domain space R^2. We want to make matrix $[T]_{B',B}$ diagonal. So we can assume that

$$[T]_{B',B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$[[T(\mathbf{e}_1)]_{B'}\ [T(\mathbf{e}_2)]_{B'}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$[T(\mathbf{e}_1)]_{B'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_{B'} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$T(\mathbf{e}_1) = 1(\mathbf{e}_1 + a\mathbf{e}_2) + 0(2\mathbf{e}_1 + b\mathbf{e}_2); \qquad T(\mathbf{e}_2) = 0(\mathbf{e}_1 + a\mathbf{e}_2) + 1(2\mathbf{e}_1 + b\mathbf{e}_2)$$
$$\mathbf{e}_1 + 5\mathbf{e}_2 = 1(\mathbf{e}_1 + a\mathbf{e}_2) + 0(2\mathbf{e}_1 + b\mathbf{e}_2); \qquad 2\mathbf{e}_1 + 4\mathbf{e}_2 = 0(\mathbf{e}_1 + a\mathbf{e}_2) + 1(2\mathbf{e}_1 + b\mathbf{e}_2)$$

Thus $\quad a = 5, b = 4$

EXAMPLE 3.34 Find the associated matrix of the following linear transformation.

Let V and W be two vector spaces of dimension m and n respectively. Suppose $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is a basis for V. The linear transformation $T: V \to W$ is defined as zero transformation by

$$T(\mathbf{v}) = \mathbf{0}, \text{ for each } \mathbf{v} \text{ in } V.$$

Solution: Since $T(\mathbf{v}) = \mathbf{0}$, for each $\mathbf{v}$ in V. Therefore,

$$T(\mathbf{v}_i) = \mathbf{0}, \qquad \text{for each } i,\ 1 \le i \le m$$

Thus the associated $m \times n$ zero matrix with the given zero transformation T is given by

$$[T] = [T(\mathbf{v}_1)\ T(\mathbf{v}_2)\ \ldots\ T(\mathbf{v}_m)]$$

$$= \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \ldots & 0 \end{bmatrix}$$

EXAMPLE 3.35 Find the associated matrix of the following linear transformation.

Let V be an n-dimensional vector space. Suppose $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V. The linear transformation $T: V \to V$ is defined as identity transformation by

$$T(\mathbf{v}) = \mathbf{v}, \text{ for each } \mathbf{v} \text{ in } V.$$

Solution: Since $T(\mathbf{v}) = \mathbf{v}$, for each $\mathbf{v}$ in V, we have

$$T(\mathbf{v}_i) = \mathbf{v}_i, \qquad \text{for each } i,\ 1 \le i \le n$$

$$T(\mathbf{v}_i) = 0(\mathbf{v}_1) + 0(\mathbf{v}_2) + \cdots + 1(\mathbf{v}_i) + \cdots + 0(\mathbf{v}_n)$$

$$T(\mathbf{v}_i) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \longrightarrow i\text{th position}; \quad \text{for each } i,\ 1 \le i \le n \\ \\ \\ \\ \end{matrix}$$

Thus the associated $n \times n$ identity matrix with the given identity transformation T is given by

$$[T] = [T(\mathbf{v}_1)\ T(\mathbf{v}_2)\ \ldots\ T(\mathbf{v}_n)]$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & . & \ldots & . \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

EXAMPLE 3.36 Find the associated matrix of the linear transformation where an orthogonal projection operator $T: R^2 \to R^2$ maps each vector into its orthogonal projection on the x-axis, that is $T(x, y) = (x, 0)$.

Solution: Consider that the basis for R^2 is a standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ for both domain and co-domain of T. Since T is defined by the formula

$$T(x, y) = (x, 0)$$

$$T(\mathbf{e}_1) = (1, 0) = 1(\mathbf{e}_1) + 0(\mathbf{e}_2); \qquad T(\mathbf{e}_2) = (0, 0) = 0(\mathbf{e}_1) + 0(\mathbf{e}_2)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_{\mathbf{B}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Therefore the associated matrix with the given linear transformation T is

$$[T]_{B,B} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

EXAMPLE 3.37 Find the associated matrix of the linear transformation where an orthogonal projection operator $T: R^3 \to R^3$ maps each vector into its orthogonal projection on the x-axis, that is, $T(x, y, \mathrm{z}) = (x, 0, 0)$.

Solution: Consider that the basis for R^3 is a standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for both domain and co-domain of T. Since T is defined by the formula,

$$T(x, y, \mathrm{z}) = (x, 0, 0).$$

$$T(\mathbf{e}_1) = (1, 0, 0) = 1(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3)$$

$$T(\mathbf{e}_2) = (0, 0, 0) = 0(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3);$$

$$T(\mathbf{e}_3) = (0, 0, 0) = 0(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_3)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

Therefore the associated matrix with the given linear transformation T is

$$[T]_{B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

EXAMPLE 3.38 Find the associated matrix of the linear transformation where the dilation and contraction operator $T: R^n \to R^n$ is defined as follows:

$$\text{Dilation: } T(\mathbf{u}) = k\mathbf{u}; \qquad k > 1$$

$$\text{Contraction: } T(\mathbf{u}) = k\mathbf{u}; \qquad 0 < k < 1$$

Solution: Consider the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for R^n. Then

$$T(\mathbf{e}_i) = k\mathbf{e}_i \qquad \text{for each } i,\ 1 \le i \le n$$

$$= 0(\mathbf{e}_1) + 0(\mathbf{e}_2) + \cdots + k(\mathbf{e}_i) + \cdots + 0(\mathbf{e}_n)$$

$$[T(\mathbf{e}_i)]_B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ k \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \longrightarrow i\text{th position}; \qquad \text{for each } i,\ 1 \le i \le n \\ \\ \\ \\ \end{matrix}$$

So the required $n \times n$ matrix for T is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B\, [T(\mathbf{e}_2)]_B\ \ldots\ [T(\mathbf{e}_n)]_B]$$

$$= \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & k \end{bmatrix}$$

EXAMPLE 3.39 Find the associated matrix of the linear transformation where the reflection operator about the x-axis in R^2 is defined as

$$T: R^2 \to R^2, \qquad T(x, y) = (x, -y)$$

Solution: For the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ of R^2,

$$T(\mathbf{e}_1) = (1, 0) = 1(\mathbf{e}_1) + 0(\mathbf{e}_2) \qquad T(\mathbf{e}_2) = (0, -1) = 0(\mathbf{e}_1) + (-1)(\mathbf{e}_2)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

So the required matrix for T is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B\, [T(\mathbf{e}_2)]_B]$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

EXAMPLE 3.40 Find the associated matrix of the linear transformation where the reflection operator about the yz-plane in R^3 is defined as

$$T: R^3 \to R^3, \qquad T(x, y, z) = (-x, y, z)$$

Solution: For the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ of R^3,

$$T(\mathbf{e}_1) = (-1, 0, 0) = (-1)(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3)$$

$$T(\mathbf{e}_2) = (0, 1, 0) = 0(\mathbf{e}_1) + 1(\mathbf{e}_2) + 0(\mathbf{e}_3)$$
$$\boldsymbol{T}(\mathbf{e}_3) = (0, 0, 1) = 0(\mathbf{e}_1) + 0(\mathbf{e}_2) + 1(\mathbf{e}_3)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_3)]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore the associated matrix of the given reflection operator is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B\ [T(\mathbf{e}_2)]_B\ [T(\mathbf{e}_3)]_B]$$
$$= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EXAMPLE 3.41 Find the associated matrix of the linear transformation where the rotation operator in R^2 rotates each vector through a fixed angle θ defined as

$$T: R^2 \to R^2, \qquad T(x, y) = (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta)$$

Solution: Consider the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ for R^2,

$$T(\mathbf{e}_1) = (\cos\theta, \sin\theta) = \cos\theta\,(\mathbf{e}_1) + \sin\theta\,(\mathbf{e}_2)$$
$$T(\mathbf{e}_2) = (-\sin\theta, \cos\theta) = (-\sin\theta)\,(\mathbf{e}_1) + (\cos\theta)\,(\mathbf{e}_2)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}; \qquad T(\mathbf{e}_2)]_B = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix};$$

So the required matrix for T is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B\ [T(\mathbf{e}_2)]_B]$$
$$= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

EXAMPLE 3.42 Find the associated matrix of the linear transformation where the rotation operator in R^3 about the x-axis with a fixed angle θ (in counterclockwise direction) is defined as

$$T: R^3 \to R^3, \qquad T(x, y, z) = (x,\ y\cos\theta - z\sin\theta,\ y\sin\theta + z\cos\theta)$$

Solution: For the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ of R^3,

$$T(\mathbf{e}_1) = (1, 0, 0) = 1(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3)$$
$$T(\mathbf{e}_2) = (0, \cos\theta, \sin\theta) = 0(\mathbf{e}_1) + \cos\theta\,(\mathbf{e}_2) + \sin\theta\,(\mathbf{e}_3)$$
$$T(\mathbf{e}_3) = (0, -\sin\theta, \cos\theta) = 0\,(\mathbf{e}_1) + (-\sin\theta)\,(\mathbf{e}_2) + (\cos\theta)\,(\mathbf{e}_3)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix}; \qquad [T(\mathbf{e}_3)]_B = \begin{bmatrix} 0 \\ -\sin\theta \\ \cos\theta \end{bmatrix}$$

Therefore the associated matrix of the given reflection operator is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B \; [T(\mathbf{e}_2)]_B \; [T(\mathbf{e}_3)]_B]$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

EXAMPLE 3.43 If $T: P_2 \to P_3$ is the linear transformation defined by $T(p(x)) = xp(x+1)$, then

(i) find the matrix $[T]_{B',B}$ with respect to the bases $B = \{1, x, x^2\}$ and $B' = \{1, x, x^2, x^3\}$.

(ii) verify the formula $[T(p(x))]_{B'} = [T]_{B',B}\,[p(x)]_B$

(iii) compute $T[1 + 2x - x^2]$ using the formula in (ii).

Solution: The given linear transformation is $T(p(x)) = xp(x+1)$,

(i) For the bases $B = \{1, x, x^2\}$ and $B' = \{1, x, x^2, x^3\}$ of P_2 and P_3 respectively,

$$\begin{aligned} T(1) &= x(1) &&= x &&= 0(1) + 1(x) + 0(x^2) + 0(x^3) \\ T(x) &= x(x+1) &&= x + x^2 &&= 0(1) + 1(x) + 1(x^2) + 0(x^3) \\ T(x^2) &= x(x+1)^2 &&= x + 2x^2 + x^3 &&= 0(1) + 1(x) + 2(x^2) + 1(x^3) \end{aligned}$$

$$[T(1)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(x)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}; \qquad [T(x^2)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

Therefore the associated matrix of T is

$$[T]_{B',B} = [[T(1)]_{B'} \; [T(x)]_{B'} \; [T(x^2)]_{B'}]$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

(ii) Consider a vector $p(x) = a_0 + a_1x + a_2x^2$ of P_2,

$$[p(x)]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

We want to verify the result $[T(p(x))]_{B'} = [T]_{B',B}\,[p(x)]_B$.

Since
$$\begin{aligned} T(p(x)) &= xp(x+1). \\ &= x[a_0 + a_1(x+1) + a_2(x+1)^2] \\ &= x[a_0 + a_1 + a_1x + a_2 + 2a_2x + a_2x^2] \\ &= x[(a_0 + a_1 + a_2) + (a_1 + 2a_2)x + a_2x^2] \end{aligned}$$

$$= 0(1) + (a_0 + a_1 + a_2)x + (a_1 + 2a_2)x^2 + a_2x^3$$

$$[T(p(x))]_{B'} = \begin{bmatrix} 0 \\ a_0 + a_1 + a_2 \\ a_1 + 2a_2 \\ a_2 \end{bmatrix}$$

$$[T]_{B',B}\,[p(x)]_B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ a_0 + a_1 + a_2 \\ a_1 + 2a_2 \\ a_2 \end{bmatrix}$$

$$= [T(p(x))]_{B'}$$

(iii) If $p(x) = 1 + 2x - x^2$, then

$$[p(x)]_B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

To compute $T[1 + 2x - x^2]$, we use the formula

$$[T(p(x))]_{B'} = [T]_{B',B}\,[p(x)]_B$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \end{bmatrix}$$

$$= 0(1) + 2(x) + 0(x^2) + (-1)x^3 = 2x - x^3$$

$$\therefore \qquad T[1 + 2x - x^2] = 2x - x^3$$

EXERCISE SET 2

Find the associated matrix of the linear transformations described below in Questions 1 to 6.

1. $T: R^2 \to R^2$, $T(x_1, x_2) = (3x_1 + 2x_2, x_1 - 4x_2)$ with bases $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ for the domain and co-domain of T.

2. $T: R^2 \to R^4$, $T(x_1, x_2) = (3x_1 - 4x_2, x_1 + 2x_2, 6x_1 - x_2, 10x_2)$ with bases $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ for the domain and co-domain of T.

3. (i) Consider the orthogonal projection operator T: $R^2 \to R^2$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y) = (0, y)$ with respect to standard basis.

(ii) Consider the orthogonal projection operator T: $R^3 \to R^3$ that maps each vector into its orthogonal projection on the y-axis, that is, $T(x, y, z) = (0, y, 0)$ with respect to standard basis.

(iii) Consider the orthogonal projection operator T: $R^3 \to R^3$ that maps each vector into its orthogonal projection on the z-axis, that is, $T(x, y, z) = (0, 0, z)$ with respect to standard basis.

4. (i) The reflection operator about y-axis in R^2 is defined as T: $R^2 \to R^2$, $T(x, y) = (-x, y)$ with respect to standard basis.

(ii) The reflection operator about xz-plane in R^3 is defined as T: $R^3 \to R^3$, $T(x, y, z) = (x, -y, z)$ with respect to standard basis.

(iii) The reflection operator about xy-plane in R^3 is defined as T: $R^3 \to R^3$, $T(x, y, z) = (x, y, -z)$ with respect to standard basis.

5. (i) In R^3, the rotation operator about y-axis with a fixed angle θ (in counterclockwise direction) is defined as

$$T: R^3 \to R^3,$$

$$T(x, y, z) = (z \sin\theta + x \cos\theta, y, z \cos\theta - x \sin\theta)$$

(ii) In R^3, the rotation operator about z-axis with a fixed angle θ (in counterclockwise direction) is defined as

$$T: R^3 \to R^3,$$

$$T(x, y, z) = (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta, z)$$

6. Let T: $R^2 \to R^2$ be a linear transformation such that

$$T(\mathbf{e}_1 + \mathbf{e}_2) = \mathbf{e}_1 - 3\mathbf{e}_2; \qquad T(2\mathbf{e}_1 - \mathbf{e}_2) = 8\mathbf{e}_1 + 15\mathbf{e}_2$$

when $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for R^2.

7. If T: $P_1 \to P_2$ is the linear transformation defined by $T(p(x)) = xp(x)$, then

(i) find the matrix $[T]_{B',B}$ with respect to the bases $B = \{1, x\}$ and $B' = \{1, x, x^2\}$.

(ii) verify the formula $[T(p(x))]_{B'} = [T]_{B',B}\,[p(x)]_B$

(iii) compute $T[1 + x]$ using the formula in (ii).

8. If T: $R^2 \to R^3$ is the linear transformation defined by $T(x, y) = (y, -5x + 13y, -7x + 16y)$, then

(i) find the matrix $[T]_{B',B}$ with respect to the bases $B = \{[3, 1], [5, 2]\}$ and $B' = \{[1,0, -1], [-1, 2, 2], [2, 1, -3]\}$.

(ii) verify the formula $[T(x, y)]_{B'} = [T]_{B',B}\,[(x, y)]_B$.

3.4 ALGEBRA OF LINEAR TRANSFORMATION

Let V and W be two vector spaces. Consider the set of all linear transformations from V to W, say

$$L(V, W) = \{T \mid T: V \to W \text{ is a linear transformation}\}$$

In this section, we will define algebraic operations like addition, scalar multiplication, composition on $L(V, W)$. We will also see that $L(V, W)$ is a vector space under these addition and scalar multiplication operations.

Definition: Addition and Scalar Multiplication of Linear Transformation

Let S: $V \to W$ and T: $V \to W$ be linear transformations. If α is any scalar, then addition $S + T$ and scalar multiplication (αT) is defined by the formula,

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}), \qquad \text{for } \mathbf{v} \in V$$

$$(\alpha T)(\mathbf{v}) = \alpha T(\mathbf{v}), \qquad \text{for all } \mathbf{v} \in V \text{ and any scalar } \alpha.$$

It is easy to show that $S + T$ and αT are also linear transformations from V to W. Hence $S + T$, $\alpha T \in L(V, W)$. The zero transformation $\mathbf{0}(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v} \in V$ and the transformation $(-T)(\mathbf{v}) = -(T(\mathbf{v}))$ are also linear. So the zero elements $\mathbf{0}$ and the additive inverse $(-T)$ are also elements of $L(V, W)$. Similarly, it is easy to verify the other axioms of vector space for $L(V, W)$. Therefore, we have the following theorem.

Theorem 3.2

The set $L(V, W)$ of all linear transformations from V into W is a vector space with operations,

$$\textit{Addition: } (T + S)(\mathbf{v}) = T(\mathbf{v}) + S(\mathbf{v}), \qquad \text{for } S, T \in L(V, W)$$

$$\textit{Scalar multiplication: } (\alpha T)(\mathbf{v}) = \alpha T(\mathbf{v}), \qquad \text{for all scalars } \alpha$$

EXAMPLE 3.44 Find $T_1 + T_2$, $2T_1 - T_2$ if T_1: $R^2 \to R^2$, $T_1(x, y) = (x + y, 2x - y)$ and T_2: $R^2 \to R^2$, $T_2(x, y) = (x, -y)$.

Solution: The given linear transformations are

$$T_1(x, y) = (x + y, 2x - y) \qquad \text{and} \qquad T_2(x, y) = (x, -y)$$

By the definitions,

$$\begin{aligned}
(T_1 + T_2)(x, y) &= T_1(x, y) + T_2(x, y) \\
&= (x + y, 2x - y) + (x, -y) \\
&= (2x + y, 2x - 2y) \\
(2T_1)(x, y) &= 2[T_1(x, y)] \\
&= 2(x + y, 2x - y) \\
&= (2x + 2y, 4x - 2y) \\
(2T_1 - T_2)(x, y) &= (2T_1)(x, y) - T_2(x, y) \\
&= (2x + 2y, 4x - 2y) - (x, -y) \\
&= (x + 2y, 4x - y)
\end{aligned}$$

Composition Operation

Definition: Composition Operation

Let V, U, W be vector spaces. If T_1: $V \to U$ and T_2: $U \to W$ are linear transformations, then composition of T_2 with T_1 is the function $T_2 \text{o} T_1$: $V \to W$ defined by the equation

$$(T_2 \text{o} T_1)(\mathbf{v}) = T_2(T_1(\mathbf{v})), \qquad \text{for all } \mathbf{v} \text{ in } V.$$

The definition means that to find the image of a vector $\mathbf{v}$ of vector space V under the mapping $T_2 \text{o} T_1$, we first find the image of $\mathbf{v}$ under the transformation T_1 and then find the image of $T_1(\mathbf{v})$ under the transformation T_2. So, one can easily observe that the range of T_1 should be contained in the domain of T_2, that is, $R(T_1) \subseteq D(T_2)$

For the vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ and some scalars α, β

$$\begin{aligned}(T_2 \circ T_1)(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) &= T_2(T_1(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2)) \\ &= T_2[\alpha T_1(\mathbf{v}_1) + \beta T_1(\mathbf{v}_2)] && \text{since } T_1 \text{ is linear} \\ &= \alpha T_2(T_1(\mathbf{v}_1)) + \beta T_2(T_1(\mathbf{v}_2)) && \text{since } T_2 \text{ is linear} \\ &= \alpha(T_2 \circ T_1)(\mathbf{v}_1) + \beta(T_2 \circ T_1)(\mathbf{v}_2)\end{aligned}$$

Thus, $$(T_2 \circ T_1)(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha(T_2 \circ T_1)(\mathbf{v}_1) + \beta(T_2 \circ T_1)(\mathbf{v}_2)).$$

The above property gives us the following theorem.

Theorem 3.3 [Composition Map]

If $T_1: V \to U$ and $T_2: U \to W$ are linear transformations, then $T_2 \circ T_1: V \to W$ is also a linear transformation.

EXAMPLE 3.45 Find the composition of the following linear transformations if it exist.

$$T_1: R^2 \to R^2, \quad T_1(x, y) = (x + y, 2x - y) \qquad \text{and} \qquad T_2: R^2 \to R^2, T_2(x, y) = (x, -y)$$

Solution: Since the range of T_1 is a subset of domain of T_2, the composition map $T_2 \circ T_1$ exists. Moreover,

$$\text{Domain of } T_2 \circ T_1 = \text{Domain of } T_1 = R^2$$

$$\text{Co-domain of } T_2 \circ T_1 = \text{Co-domain of } T_2 = R^2$$

Thus, $$T_2 \circ T_1: R^2 \to R^2$$

and $$\begin{aligned}(T_2 \circ T_1)(x, y) &= T_2(T_1(x, y)) \\ &= T_2(x + y, 2x - y) && \text{(by definition of } T_1) \\ &= (x + y, -(2x - y)) && \text{(by definition of } T_2) \\ &= (x + y, y - 2x)\end{aligned}$$

The other composition $T_1 \circ T_2$ exists because the range of T_2 is contained in R^2, a domain of T_1. Moreover,

$$\text{Domain of } T_1 \circ T_2 = \text{Domain of } T_2 = R^2$$

$$\text{Co-domain of } T_1 \circ T_2 = \text{Co-domain of } T_1 = R^2$$

Thus, $$T_1 \circ T_2: R^2 \to R^2$$

$$\begin{aligned}(T_1 \circ T_2)(x, y) &= T_1(T_2(x, y)) \\ &= T_1(x, -y) && \text{(by definition of } T_2) \\ &= (x - y, 2x + y) && \text{(by definition of } T_1)\end{aligned}$$

EXAMPLE 3.46 Find the composition of the following linear transformation if it exists.

$$T_1: R^2 \to R^3, T_1(x, y) = (x - y, 2y, x + 3y)$$

$$T_2: R^3 \to R^2, T_2(x, y, z) = (x + y + z, x - z)$$

Solution: Since the range of T_1 is a subset of a domain of T_2, the composition map $T_2 \circ T_1$ exists. Moreover,

$$\text{Domain of } T_2 \circ T_1 = \text{Domain of } T_1 = R^2$$

$$\text{Co-domain of } T_2 \circ T_1 = \text{Co-domain of } T_2 = R^2$$

Thus, $T_2 \circ T_1 : R^2 \to R^2$

and
$$\begin{aligned}(T_2 \circ T_1)(x, y) &= T_2(T_1(x, y)) \\ &= T_2(x - y, 2y, x + 3y) && \text{(by definition of } T_1) \\ &= (2x + 4y, -4y) && \text{(by definition of } T_2)\end{aligned}$$

The other composition $T_1 \circ T_2$ exists because the range of T_2 is contained in R^2, a domain of T_1. Moreover,

$$\text{Domain of } T_1 \circ T_2 = \text{Domain of } T_2 = R^3$$
$$\text{Co-domain of } T_1 \circ T_2 = \text{Co-domain of } T_1 = R^3$$

Thus, $T_1 \circ T_2 : R^3 \to R^3$

and
$$\begin{aligned}(T_1 \circ T_2)(x, y, z) &= T_1(T_2(x, y, z)) \\ &= T_1(x + y + z, x - z) && \text{(by definition of } T_2) \\ &= (y + 2z, 2x - 2z, 4x + y - 2z) && \text{(by definition of } T_1)\end{aligned}$$

EXAMPLE 3.47 Find the composition of the following linear transformation if it exists.

$$T_1 : R^2 \to R^3,\ T_1(x, y) = (x, y, x + y)$$
$$T_2 : R^3 \to R^3,\ T_2(x, y, z) = (x, x + y, x + y + z)$$

Solution: Since the range of T_1 is a subset of a domain of T_2, the composition map $T_2 \circ T_1$ exists. Moreover,

$$\text{Domain of } T_2 \circ T_1 = \text{Domain of } T_1 = R^2$$
$$\text{Co-domain of } T_2 \circ T_1 = \text{Co-domain of } T_2 = R^3$$

Thus, $T_2 \circ T_1 : R^2 \to R^3$

and
$$\begin{aligned}(T_2 \circ T_1)(x, y) &= T_2(T_1(x, y)) \\ &= T_2(x, y, x + y) && \text{(by definition of } T_1) \\ &= (x, x + y, 2x + 2y) && \text{(by definition of } T_2)\end{aligned}$$

The other composition $T_1 \circ T_2$ does not exist because the range of T_2 is a subset of R^3 and the domain of T_1 is R^2 which cannot contain a subset of R^3.

EXAMPLE 3.48 Find the composition of the following linear transformation if it exists.

$$T_1 : M_{22} \to R,\ T_1\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc;$$

$$T_2 : M_{22} \to M_{22},\ T_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Solution: The composition map $T_2 \circ T_1$ does not exist because the range of T_1 is R which is not contained in domain M_{22} of T_2. But the range of T_2 is a subset of M_{22} which is a domain of T_1. Therefore the composition map $T_1 \circ T_2$ exists. Moreover,

$$\text{Domain of } T_1 \circ T_2 = \text{Domain of } T_2 = M_{22}$$
$$\text{Co-domain of } T_1 \circ T_2 = \text{Co-domain of } T_1 = R$$

Thus, $T_1 \text{o} T_2 \colon M_{22} \to R$

and
$$(T_1 \text{o} T_2)\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = T_1\left(T_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\right)$$

$$= T_1\left(\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) \qquad \text{(by definition of } T_2)$$

$$= ad - bc \qquad \text{(by definition of } T_1)$$

EXAMPLE 3.49 Find the composition of the following linear transformation if it exists.

$$T_1 \colon P_2 \to P_3,\ T_1(p(x)) = xp(x);\quad T_2 \colon P_3 \to P_2,\ T_2(p(x)) = p'(x)$$

Solution: Since the range of T_1 is a subset of a domain of T_2, the composition map $T_2 \text{o} T_1$ exists. Moreover,

$$\text{Domain of } T_2\text{o}T_1 = \text{Domain of } T_1 = P_2$$

$$\text{Co-domain of } T_2\text{o}T_1 = \text{Co-domain of } T_2 = P_2$$

Thus, $T_2\text{o}T_1 \colon P_2 \to P_2$

and
$$(T_2\text{o}T_1)\,(p(x)) = T_2(T_1(p(x)))$$

$$= T_2(x\,p(x)) \qquad \text{(by definition of } T_1)$$

$$= p(x) + xp'(x) \qquad \text{(by definition of } T_2)$$

Since the range of T_2 is a subset of P_2 which is the domain of T_1, the composition map $T_1\text{o}T_2$ exists. Moreover,

$$\text{Domain of } T_1\text{o}T_2 = \text{Domain of } T_2 = P_3$$

$$\text{Co-domain of } T_1\text{o}T_2 = \text{Co-domain of } T_1 = P_3$$

Thus, $T_1\text{o}T_2 \colon P_3 \to P_3$

and
$$(T_1\text{o}T_2)\,(p(x)) = T_1(T_2(p(x)))$$

$$= T_1(p'(x)) \qquad \text{(by definition of } T_2)$$

$$= xp'(x) \qquad \text{(by definition of } T_1)$$

EXAMPLE 3.50 If $T_1(x) = (x, -x)$, $T_2(x, y) = (x, y, x + y)$, and $T_3(x, y, z) = (z, x, y)$, then find $T_3\text{o}T_2\text{o}T_1$.

Solution: The given linear transformations are

$$T_1 \colon R \to R^2, \qquad T_1(x) = (x, -x)$$

$$T_2 \colon R^2 \to R^3, \qquad T_2(x, y) = (x, y, x + y)$$

$$T_3 \colon R^3 \to R^3, \qquad T_3(x, y, z) = (z, x, y)$$

By the definition of the composition map

$$(T_3\text{o}T_2\text{o}T_1)\,(x) = T_3(T_2\,(T_1(x)))$$

$$= T_3(T_2(x, -x))$$

$$= T_3(x, -x, 0)$$

$$= (0, x, -x)$$

EXAMPLE 3.51 If $T: R^2 \to R^2$ is a reflection of R^2 about the x-axis, then show that $ToT = \mathbf{I}$.

Solution: Here $T: R^2 \to R^2$ is a reflection of R^2 about the x-axis,

$$T(x, y) = (x, -y)$$

By the definition of composition map, the $ToT: R^2 \to R^2$ is defined as

$$\begin{aligned}(ToT)(x, y) &= T(T(x, y)) \\ &= T(x, -y) \\ &= (x, y) \\ &= \mathbf{I}(x, y)\end{aligned}$$

Hence $\quad ToT = \mathbf{I}$.

As composition map T_2oT_1 is a linear transformation. So we can associate matrix with T_2oT_1. The following theorem gives the formula of matrix of T_2oT_1.

Theorem 3.4 [Matrix of Composition Map]

(i) Let $T_1: V \to U$ and $T_2: U \to W$ be two linear transformations. If B, B'' and B' are bases for V, U and W respectively, then $[T_2oT_1]_{B', B} = [T_2]_{B', B''}[T_1]_{B'', B}$.

(ii) Let $T_1: V \to U$ and $T_2: U \to W$ and $T_3: W \to X$ be linear transformations. If B, B'', B''' and B' are bases for the vector space V, U, W and X respectively, then

$$[T_3oT_2oT_1]_{B', B} = [T_3]_{B', B'''}[T_2]_{B''', B''}[T_1]_{B'', B}.$$

EXAMPLE 3.52 Find $[T_2oT_1]_{B, B}$ for the following linear transformations.

$$T_1: R^2 \to R^2,\ T_1(x, y) = (x + y, 2x - y)$$

and $\quad T_2: R^2 \to R^2,\ T_2(x, y) = (x, -y)$

with the basis $B = \{[1, 0], [0, 1]\}$ for R^2. Also, verify the result of Theorem 3.4.

Solution: For the linear transformation,

$$T_1(x, y) = (x + y, 2x - y)$$
$$T_1(1, 0) = (1, 2) = 1(1, 0) + 2(0, 1)$$
$$T_1(0, 1) = (1, -1) = 1(1, 0) + (-1)(0, 1)$$
$$[T_1]_{B,B} = [[T_1(\mathbf{e}_1)]_B\ [T_1(\mathbf{e}_2)]_B] = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

For the linear transformation,

$$T_2(x, y) = (x, -y)$$
$$T_2(1, 0) = (1, 0) = 1(1, 0) + 0(0, 1)$$
$$T_2(0, 1) = (0, -1) = 0(1, 0) + (-1)(0, 1)$$
$$[T_2]_{B,B} = [[T_2(\mathbf{e}_1)]_B\ [T_2(\mathbf{e}_2)]_B] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$[T_2oT_1]_{B,B} = [T_2]_{B,B}[T_1]_{B,B}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$$

Look at Example 3.45,

$$(T_2 o T_1)\,(x) = (x + y, y - 2x)$$

$$(T_2 o T_1)\,(1, 0) = (1, -2) = 1(1, 0) + (-2)\,(0, 1)$$

$$(T_2 o T_1)\,(0, 1) = (1, 1) = 1(1, 0) + (1)\,(0, 1)$$

$$[T_2 o T_1]_{B,B} = [[T_2 o T_1\,(1, 0)]_B\,[T_2 o T_1\,(0, 1)_B]]$$

$$= \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$$

From Theorem 3.4 (i) and (ii),

$$[T_2 o T_1]_{B,B} = [T_2]_{B,B}\,[T_1]_{B,B}$$

EXAMPLE 3.53 Find $[T_2 o T_1]_{B,\,B'}$ for the following linear transformations.

$$T_1 : P_2 \to P_3,\; T_1(p(x)) = x\,p(x)$$

and $$T_2 : P_3 \to P_2,\; T_2(p(x)) = p'(x)$$

with the basis $B = \{1, x, x^2\}$ for P_2 and $B' = \{1, x, x^2, x^3\}$ for P_3. Also, verify the result of Theorem 3.4.

Solution: For the linear transformation,

$$T_1(p(x)) = xp(x)$$

$$T_1(1) = x; \qquad T_1(x) = x^2; \qquad T_1(x^2) = x^3$$

$$T_1(1) = 0(1) + 1(x) + 0(x^2) + 0(x^3)$$

$$T_1(x) = 0(1) + 0(x) + 1(x^2) + 0(x^3)$$

$$T_1(x^2) = 0(1) + 0(x) + 0(x^2) + 1(x^3)$$

$$[T_1(1)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T_1(x)]_{B'} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; \qquad [T_1(x^2)]_{B'} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$[T_1]_{B',B} = [[T_1(1)]_{B'} \;\; [T_1(x)]_{B'} \;\; [T_1(x^2)]_{B'}]$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

For the linear transformation,

$$T_2(p(x)) = p'(x)$$

$$T_2(1) = 0 \quad = 0(1) + 0(x) + 0(x^2)$$

$$T_2(x) = 1 \quad = 1(1) + 0(x) + 0(x^2)$$

$$T_2(x^2) = 2x \;= 0(1) + 2(x) + 0(x^2)$$

$$T_2(x^3) = 3x^2 = 0(1) + 0(x) + 3(x^2)$$

$$[T_2(1)]_B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; \quad \left[T_2(x)\right]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad \left[T_2(x^2)\right]_{B'} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}; \quad \left[T_2(x^3)\right]_{B'} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

$$[T_2]_{B,B'} = [[T_2(1)]_{B'}\, [T_2(x)]_{B'}\, [T_2(x^2)]_{B'}\, [T_2(x^3)]_{B'}]$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

$$[T_2]_{B,B'}\, [T_1]_{B',B} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \tag{i}$$

By the definition of composition.

$$(T_2 o T_1)\,(p(x)) = (xp(x))' = p(x) + xp'(x)$$

$$(T_2 o T_1)\,(1) = 1 = 1 + 0(x) + 0(x^2)$$

$$(T_2 o T_1)\,(x) = 0 + 2(x) + 0(x^2)$$

$$(T_2 o T_1)\,(x^2) = 0 + 0(x) + 3x^2$$

Thus
$$[T_2 o T_1]_{B,\,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \tag{ii}$$

From (i) and (ii),

$$[T_2 o T_1]_{B,B} = [T_2]_{B,B'}\, [T_1]_{B',B}$$

EXAMPLE 3.54 Use the formula of Theorem 3.4, to find the matrix of $T_3 o T_2 o T_1$ if $T_1(x) = (x, -x)$, $T_2(x, y) = (x, y, x + y)$, $T_3(x, y, z) = (z, x, y)$ and $B = \{1\}$, $B'' = \{[1, 0], [0, 1]\}$ and $B' = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ are the basis for R, R^2 and R^3 respectively.

Solution: For $T_1(x) = (x, -x)$

$$T_1(1) = (1, -1) = 1(1, 0) + (-1)(1, 0)$$

$$[T_1]_{B'',B} = [[T_1(1)]_{B''}] = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

For
$$T_2(x, y) = (x, y, x + y)$$
$$T_2(1, 0) = (1, 0, 1) = 1(1, 0, 0) + 0(0, 1, 0) + 1(0, 0, 1)$$
$$T_2(0, 1) = (0, 1, 1) = 0(1, 0, 0) + 1(0, 1, 0) + 1(0, 0, 1);$$
$$[T_2(1, 0)]_{B'} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}; \qquad [T_2(0,1)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$
$$[T_2]_{B',B''} = [[T_2(\mathbf{e}_1)]_{B'} \quad [T_2(\mathbf{e}_2)]_{B'}]$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

For
$$T_3(x, y, z) = (z, x, y)$$
$$T_3(1, 0, 0) = (0, 1, 0) = 0(1, 0, 0) + 1(0, 1, 0) + 0(0, 0, 1)$$
$$T_3(0, 1, 0) = (0, 0, 1) = 0(1, 0, 0) + 0(0, 1, 0) + 1(0, 0, 1)$$
$$T_3(0, 0, 1) = (1, 0, 0) = 1(1, 0, 0) + 0(0, 1, 0) + 0(0, 0, 1)$$
$$[T_3(1, 0, 0)]_{B'} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad [T_3(0,1,0)]_{B'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; \qquad [T_3(0,1,0)]_{B'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix};$$
$$[T_3]_{B',B'} = [[T_3(1, 0, 0)]_{B'} [T_3(0, 1, 0)]_{B'} [T_3(0, 0, 1)]_{B'}$$
$$= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

By Theorem 3.4,
$$[T_3 o T_2 o T_1]_{B',B} = [T_3]_{B',B'} [T_2]_{B',B''} [T_1]_{B'',B}$$
$$= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

EXERCISE SET 3

1. Find $T_1 + T_2$, $5T_1$ if

$$T_1: R^3 \to R^2,\ T_1(x, y, z) = (x - 2y, y + 3z)$$

and

$$T_2: R^3 \to R^2,\ T_2(x, y, z) = (3x, y - z).$$

2. Find the composition functions $T_1 o T_2$, $T_2 o T_1$ of the following linear transformations if they exist.

$$T_1: R^2 \to R^2,\ T_1(x, y) = (3x - y, x + y)$$

and

$$T_2: R^2 \to R^2,\ T_2(x, y) = (x + 2y, x - y).$$

3. If $T_1(x) = (x, 0)$, $T_2(x, y) = (x + y, x - y, y)$, and $T_3(x, y, z) = (-x, -y, -z)$, then find $T_3 o T_2 o T_1$.

4. If the linear operators $T_1: P_3 \to P_3$, $T_1(p(x)) = xp'(x)$, and $T_2: P_3 \to P_3$, $T_2(q(x)) = q(1 + x)$, verify $[T_2 o T_1]_{B,B} = [T_2]_{B,B}\,[T_1]_{B,B}$ with respect to the basis $B = \{1, x, x^2, x^3\}$

5. Find $[ToT]_{B,B}$ for the following linear transformation,

$$T: R^2 \to R^2,\ T(x, y) = (2x + y, x + 3y)$$

with respect to the basis $B = \{[1, 0], [1, 1]\}$ for R^2.

6. If $T_1: R^2 \to R^2$ and $T_2: R^2 \to R^2$ are the linear operators that rotate vectors through the angles θ_1 and θ_2 respectively, then find the matrix of the composition $[T_2 o T_1]$ with respect to standard basis.

3.5 ONE-TO-ONE AND INVERSE LINEAR TRANSFORMATIONS

Let V and W be two vector spaces with the same scalars. Consider $T: V \to W$ a linear transformation, we will define one-to-one property and inverse of linear transformation. We will also discuss how they are related with each other.

Definition: One-to-one Linear Transformation

If the linear transformation $T: V \to W$ maps distinct vectors in V into distinct vectors in W, then T is called one-to-one linear transformation.

Remark: It follows from the definition that for each vector $\mathbf{w}$ in the range of one-to-one transformation T, there is exactly one vector $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{w}$, that is, if $T(\mathbf{x}_1) = T(\mathbf{x}_2)$, then $\mathbf{x}_1 = \mathbf{x}_2$.

EXAMPLE 3.55 Determine whether T is one-to-one linear transformation:

(i) $T: R^2 \to R^2$, $T(x, y) = (x + y, 2x - y)$

(ii) $T: R^2 \to R^3$, $T(x, y) = (x - y, 2y, x + 3y)$

(iii) $T: R^3 \to R^2$, $T(x, y, z) = (x + y + z, x - z)$

(iv) $T: P_3 \to P_2$, $T(p(x)) = p'(x)$

Solution:

(i) To check the one-to-one linear transformation of T, we have to show that the images of distinct vectors of R^2 (domain) under T are same as the vectors.

Let $\mathbf{u} = [x_1, y_1]$, $\mathbf{v} = [x_2, y_2]$ be two vectors in R^2.

Suppose $T(\mathbf{u}) = T(\mathbf{v})$

$$T(x_1, y_1) = T(x_2, y_2)$$

$$(x_1 + y_1, 2x_1 - y_1) = (x_2 + y_2, 2x_2 - y_2)$$

$$x_1 + y_1 = x_2 + y_2\,; \qquad 2x_1 - y_1 = 2x_2 - y_2$$

$$x_1 = x_2; \qquad y_1 = y_2$$

$$\therefore \qquad [x_1, y_1] = [x_2, y_2]$$

Hence $\mathbf{u} = \mathbf{v}$

Therefore T is a one-to-one map.

(ii) Let $\mathbf{u} = [x_1, y_1]$, $\mathbf{v} = [x_2, y_2]$ be two vectors of R^2.

Suppose $T(\mathbf{u}) = T(\mathbf{v})$

$$T(x_1, y_1) = T(x_2, y_2)$$

$$(x_1 - y_1,\ 2y_1,\ x_1 + 3y_1) = (x_2 - y_2,\ 2y_2,\ x_2 + 3y_2)$$

$$x_1 - y_1 = x_2 - y_2; \quad 2y_1 = 2y_2; \quad x_1 + 3y_1 = x_2 + 3y_2$$

$$x_1 = x_2; \quad y_1 = y_2$$

$$\therefore \quad [x_1, y_1] = [x_2, y_2]$$

Hence $\mathbf{u} = \mathbf{v}$

Therefore, T is a one-to-one map.

(iii) Let $\mathbf{u} = [x_1, y_1, z_1]$, $\mathbf{v} = [x_2, y_2, z_2]$ be two vectors of R^3.

Suppose $T(\mathbf{u}) = T(\mathbf{v})$

$$T(x_1, y_1, z_1) = T(x_2, y_2, z_2)$$

$$(x_1 + y_1 + z_1, x_1 - z_1) = (x_2 + y_2 + z_2, x_2 - z_2)$$

$$x_1 + y_1 + z_1 = x_2 + y_2 + z_2; \quad x_1 - z_1 = x_2 - z_2$$

$$(x_1 - x_2) + (y_1 - y_2) + (z_1 - z_2) = 0; \quad (x_1 - x_2) - (z_1 - z_2) = 0$$

$$x' + y' + z' = 0; \quad x' - z' = 0 \quad \text{where } x' = x_1 - x_2,\ y' = y_1 - y_2,\ z' = z_1 - z_2$$

It is a system of linear equations with three variables and two equations. So, it has infinitely many solutions. If we consider $x = z = k$ then $y = -2k$, that is, for each other value of k, the system has one solution. If we take $k = 1$, then $\mathbf{u} = [1, -2, 1]$ and if $k = 2$, then $\mathbf{v} = [2, -4, 2]$ are solutions of the system.

Since $\mathbf{u} \neq \mathbf{v}$, but $T(\mathbf{u}) = T(\mathbf{v}) = (0, 0)$. Hence the given linear transformation is not a one-to-one linear transformation.

Remark: The above example suggests that if $\dim V = \dim R^3 = 3 > \dim W = \dim R^2 = 2$, then the linear transformation $T: V \to W$ cannot be one-to-one.

(iv) Consider the vectors $\mathbf{u} = x^2$, $\mathbf{v} = 1 + x^2$ of P_3.

Since $\mathbf{u} \neq \mathbf{v}$, but $T(\mathbf{u}) = T(\mathbf{v}) = 2x$. Therefore, T is not a one-to-one linear transformation. Moreover $\dim V = \dim P_3 = 4 > \dim W = \dim P_2 = 3$. Hence T is not a one-to-one map.

Theorem 3.5 [Linear Transformation]

If $T: V \to W$ be a linear transformation and $\dim V = n$, then the following statements are equivalent.

(i) T is one-to-one on V.

(ii) If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are independent elements in V, then $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ are independent elements in $T(V)$.

(iii) If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for V, then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)\}$ is a basis for $T(V)$.

(iv) $\dim T(V) = n$.

Inverse of Linear Transformation

Definition: Inverse of Linear Transformation

Let $T: V \to W$ be a one-to-one linear transformation, then the inverse of T is a map

$T^{-1}: T(V) \to V$ such that $T^{-1}(\mathbf{w}) = \mathbf{v}$ if $T(\mathbf{v}) = \mathbf{w}$ for all $\mathbf{v}$ in V and $\mathbf{w}$ in $T(V)$.

Moreover, T has the following property:

$$(T^{-1}oT)\,(\mathbf{v}) = T^{-1}(T(\mathbf{v})) = T^{-1}(\mathbf{w}) = \mathbf{v} = \mathbf{I}(\mathbf{v}) \qquad \text{since } \mathbf{I}: V \to V$$

$$(ToT^{-1})(\mathbf{w}) = T(T^{-1}(\mathbf{w})) = T(\mathbf{v}) = \mathbf{w} = \mathbf{I}(\mathbf{w}) \qquad \text{since } \mathbf{I}: W \to W$$

That is $\quad T^{-1}oT = I$ and $ToT^{-1} = I$.

Remark: As seen earlier, every Euclidean linear operator $T: R^n \to R^n$ can be expressed by the formula $T(\mathbf{x}) = \mathbf{Ax}$ where $\mathbf{A}$ is an $n \times n$ matrix. Therefore, the definition of inverse of such linear transformation is reduced as follows:

$$T^{-1}(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x} \qquad \text{if it exists.}$$

Therefore, we have the following theorem for the Euclidean linear operator.

Theorem 3.6 [*For Euclidean Linear Operator*]

If $T: R^n \to R^n$ is a linear transformation defined by the formula $T(\mathbf{x}) = \mathbf{Ax}$, then the following statements are equivalent.

(i) T is one-to-one
(ii) The range of T is R^n
(iii) A is invertible
(iv) T is invertible.

EXAMPLE 3.56 Find the inverse of the following linear transformations if they exist.

(i) $T: R^2 \to R^2$, $T(x, y) = (x + y, 2x - y)$
(ii) $T: R^2 \to R^3$, $T(x, y) = (x - y, 2y, x + 3y)$
(iii) $T: R^3 \to R^2$, $T(x, y, z) = (x + y + z, x - z)$
(iv) $T: P_2 \to P_2$, $T(p(x)) = p(x + 1)$
(v) $T: P_3 \to P_2$, $T(p(x)) = p'(x)$

Solution:

(i) By the definition of inverse of T, it exists if T is a one-to-one transformation. In Example 3.55(i), we have already proved that T is a one-to-one map. Therefore, its inverse map T^{-1} exists.

Let $\mathbf{w} = [w_1, w_2]$ be a vector of $T(R^2)$. We want to find $\mathbf{v} = [x_1, x_2]$ a vector in R^2 such that

$$T(\mathbf{v}) = \mathbf{w}$$

$$T(x_1, x_2) = (w_1, w_2)$$

$$(x_1 + x_2, 2x_1 - x_2) = (w_1, w_2)$$

$$x_1 + x_2 = w_1, \quad 2x_1 - x_2 = w_2$$

$$x_1 = \frac{w_1 + w_2}{3}; \qquad x_2 = \frac{2w_1 - w_2}{3}$$

$$\therefore \qquad \mathbf{v} = [x_1, x_2] = \left[\frac{w_1 + w_2}{3}, \frac{2w_1 - w_2}{3}\right]$$

Hence, the inverse of T is defined by the formula

$$T^{-1}(\mathbf{w}) = \mathbf{v}$$

$$T^{-1}(\mathbf{w}) = \left[\frac{w_1 + w_2}{3}, \frac{2w_1 - w_2}{3}\right]$$

Hence, $$T^{-1}(1, 0) = \left(\frac{1}{3}, \frac{2}{3}\right); \qquad T^{-1}(0, 1) = \left(\frac{1}{3}, -\frac{1}{3}\right)$$

Another method

Since T: $R^2 \to R^2$ is a Euclidean linear operator, we can use Theorem 3.6 to find the inverse of T. First, we have to find the matrix of T.

Here $$T(x, y) = (x + y, 2x - y)$$

$$T(1, 0) = (1, 2), \quad T(0, 1) = (1, -1)$$

$$\text{Matrix of } T = \mathbf{A} = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)] = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

$$\det \mathbf{A} = -3 \neq 0$$

So $\mathbf{A}^{-1}$ exists. Hence T^{-1} exists because $\mathbf{A}^{-1}$ exists.

By the definition,

$$T^{-1}(\mathbf{w}) = \mathbf{A}^{-1}\mathbf{w} \qquad \text{for all } \mathbf{w} \text{ in } T(R)$$

$$T^{-1}(\mathbf{w}) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{w_1 + w_2}{3} \\ \frac{2w_1 - w_2}{3} \end{bmatrix}$$

Hence $$T^{-1}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}; \qquad T^{-1}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

(ii) See Example 3.55(ii), T is a one-to-one transformation. So the inverse of T exists.
Let $\mathbf{w} = [w_1, w_2, w_3]$ be a vector of $T(R^2) \subseteq R^3$.
We want to find $\mathbf{v} = [x_1, x_2]$ in R^2 such that

$$T(\mathbf{v}) = \mathbf{w}$$

$$(x_1 - x_2, 2x_2, x_1 + 3x_2) = (w_1, w_2, w_3)$$

$$x_1 - x_2 = w_1, \quad 2x_2 = w_2, \quad x_1 + 3x_2 = w_3$$

$$x_1 = \frac{2w_1 + w_2}{2}; \quad x_2 = \frac{w_2}{2}$$

$$\mathbf{v} = [x_1, x_2] = \left[\frac{2w_1 + w_2}{2}, \frac{w_2}{2}\right]$$

Therefore, the inverse of T is given by the formula

$$T^{-1}(\mathbf{w}) = \mathbf{v}$$

$$T^{-1}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) = \left[\frac{2w_1 + w_2}{2}, \frac{w_2}{2}\right]$$

(iii) See Example 3.55(iii), T is not a one-to-one map. Therefore, the inverse of T does not exist.

(iv) To check the existence of inverse, we have to check if T is a one-to-one linear transformation.

Let $\mathbf{p}(x) = p_0 + p_1x + p_2x^2$, $\mathbf{q}(x) = q_0 + q_1x + q_2x^2$ be two vectors of P_2

Suppose $T(p(x)) = T(q(x))$

$$\mathbf{p}(x + 1) = \mathbf{q}(x + 1)$$

$$p_0 + p_1(x + 1) + p_2(x + 1)^2 = q_0 + q_1(x + 1) + q_2(x + 1)^2$$

$$p_0 = q_0,\ p_1 = q_1,\ p_2 = q_2$$

$$\therefore \qquad \mathbf{p}(x) = \mathbf{q}(x)$$

Hence T is a one-to-one transformation. So the inverse of T exists.

Let $\mathbf{w}(x) = w_0 + w_1x + w_2x^2$ be a vector of P_2 (co-domain space).

We want to find $\mathbf{p}(x) = p_0 + p_1x + p_2x^2$ in P_2 (domain) such that,

$$T(\mathbf{p}(x)) = \mathbf{w}(x)$$

$$\mathbf{p}(x + 1) = \mathbf{w}(x)$$

$$p_0 + p_1(x + 1) + p_2(x + 1)^2 = w_0 + w_1x + w_2x^2$$

$$(p_0 + p_1 + p_2) + (p_1 + 2p_2)x + p_2x^2 = w_0 + w_1x + w_2x^2$$

$$p_0 + p_1 + p_2 = w_0, \qquad p_1 + 2p_2 = w_1, p_2 = w_2$$

$$p_2 = w_2,\ p_1 = w_1 - 2w_2,\ p_0 = w_0 - w_1 + w_2$$

$$\begin{aligned} \mathbf{p}(x) &= p_0 + p_1x + p_2x^2 \\ &= (w_0 - w_1 + w_2) + (w_1 - 2w_2)x + (w_2)x^2 \\ &= w_0 + w_1(x - 1) + w_2(x - 1)^2 \\ &= \mathbf{w}(x - 1) \end{aligned}$$

Therefore, the inverse of T is given by the formula.

$$T^{-1}(\mathbf{w}(x)) = \mathbf{p}(x)$$

$$T^{-1}(\mathbf{w}(x)) = \mathbf{w}(x - 1)$$

(v) See Example 3.55(iv), T is not a one-to-one map. So the inverse of T does not exist.

Theorem 3.7 [Composition Map]

Let, T_1: $V \to U$, T_2: $U \to W$ be two one-to-one linear transformations, then

(i) $T_2 o T_1$ is one-to-one

(ii) $(T_2 o T_1)^{-1} = T_1^{-1} o T_2^{-1}$.

EXAMPLE 3.57 Verify the result of Theorem 3.7 for T_1: $R^2 \to R^2$, $T_1(x, y) = (x + y, 2x - y)$ and T_2: $R^2 \to R^2$, $T_2(x, y) = (x, -y)$.

Solution: The composition map $T_2 o T_1$: $R^2 \to R^2$ is given by the formula

$$T_2 o T_1(x, y) = (x + y, y - 2x)$$

Let $\mathbf{u} = [x_1, y_1]$, $\mathbf{v} = [x_2, y_2]$ be two vectors of domain space R^2.

Suppose $T_2 o T_1(\mathbf{u}) = T_2 o T_1(\mathbf{v})$

$$T_2 o T_1(x_1, y_1) = T_2 o T_1(x_2, y_2)$$

$$(x_1 + y_1, y_1 - 2x_1) = (x_2 + y_2, y_2 - 2x_2)$$

$$x_1 + y_1 = x_2 + y_2; \qquad y_1 - 2x_1 = y_2 - 2x_2$$

$$x_1 = x_2, \quad y_1 = y_2$$

$$\therefore \qquad \mathbf{u} = \mathbf{v}$$

Thus, $T_2 o T_1$ is a one-to-one linear transformation. Therefore its inverse $(T_2 o T_1)^{-1}$ exists.

Let $\mathbf{w} = [w_1, w_2]$ be a vector of domain space R^2.

We want to find $\mathbf{v} = [x, y]$ such that

$$(T_2 o T_1)(\mathbf{v}) = \mathbf{w}$$

$$(T_2 o T_1)(x, y) = (w_1, w_2)$$

$$(x + y, y - 2x) = (w_1, w_2)$$

$$x + y = w_1, \qquad y - 2x = w_2$$

$$x = \frac{w_1 - w_2}{3}; \quad y = \frac{2w_1 + w_2}{3}$$

$$\therefore \qquad \mathbf{v} = [x, y] = \left[\frac{w_1 - w_2}{3}, \frac{2w_1 + w_2}{3}\right]$$

Therefore, the inverse of $T_2 o T_1$ is given by the formula

$$(T_2 o T_1)^{-1}(\mathbf{w}) = \mathbf{v}$$

$$(T_2 o T_1)^{-1}(w_1, w_2) = \left(\frac{w_1 - w_2}{3}, \frac{2w_1 + w_2}{3}\right)$$

Similarly the formulas for T_1^{-1} and T_2^{-1} are as follows:

$$T_1^{-1}(w_1, w_2) = \left(\frac{w_1 + w_2}{3}, \frac{2w_1 - w_2}{3}\right)$$

$$T_2^{-1}(w_1, w_2) = (w_1, -w_2)$$

$$\begin{aligned}(T_1^{-1} o T_2^{-1})(w_1, w_2) &= T_1^{-1}(T_2^{-1}(w_1, w_2)) \\ &= T_1^{-1}(w_1, -w_2) \\ &= \left(\frac{w_1 - w_2}{3}, \frac{2w_1 + w_2}{3}\right)\end{aligned}$$

From Theorem 3.7(i) and (ii)

$$(T_2 o T_1)^{-1} = T_1^{-1} o T_2^{-1}.$$

EXERCISE SET 4

1. Determine whether T is a one-to-one linear transformation

(i) $T: R^2 \to R^2$, $T(x, y) = (10y, -x)$

(ii) $T: R^3 \to R^3$, $T(x, y, z) = (x + 1, y + 2, z + 3)$

(iii) $T: R^3 \to P_2$, $T(\alpha, \beta, \gamma) = \alpha + (\alpha + \beta)x + (\alpha + \gamma)x^2$.

(iv) Let $G = \{c_1 \cos\theta + c_2 \sin\theta \mid c_1, c_2\, \varepsilon R\}$. $T: G \to R^2$, $T(c_1\cos\theta + c_2\sin\theta) = (c_1, c_2)$.

(v) $T: M_{22} \to R$, $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

(vi) $T: P_2 \to P_2$, $T(p(x)) = p(x - 1)$.

2. Find the inverse of the following linear transformations if they exist.

(i) $T: R^2 \to R^2$, $T(x, y) = (5x + 2y, 2x + y)$

(ii) $T: R^3 \to R^3$, $T(x, y, z) = (x - y + z, 2y - z, 2x + 3y)$

(iii) $T: R^3 \to R^3$, $T(x, y, z) = (x + 5y + 2z, x + 2y + z, -x + y)$

(iv) $T: R^3 \to R^3$, $T(\mathbf{x}) = \mathbf{Ax}$ where $\mathbf{A} = \begin{bmatrix} 1 & 4 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$

3. Let $T_1: P_2 \to P_3$, $T_1(p(x)) = xp(x)$ and $T_2: P_3 \to P_3$, $T(p(x)) = p(x + 1)$ be linear transformations. Is $T_2 o T_1$ a one-to-one? If yes, then verify $(T_2 o T_1)^{-1} = T_1^{-1} o T_2^{-1}$.

3.6 KERNEL AND RANGE

In this section, we will derive two special sets, *kernel* and *range* of the linear transformation. We will also see some properties of these sets.

Definition: **Kernal and Range of Linear Transformation**

Let $T: V \to W$ be a linear transformation. Then the set of all vectors in V that T maps into $\mathbf{0}$ is called the *kernel* of T, denoted by ker (T), and the set of all images of vectors of V under T is called the *range* of T, denoted by R(T). In other words,

$$\ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\} \subseteq V$$

$$R(T) = \{T(\mathbf{v}) \in W \mid \mathbf{v} \in V\} \subseteq W$$

Theorem 3.8 [Kernal and Range of Linear Transformation]

If $T : V \to W$ is a linear transformation, then

(i) the kernel of T is a subspace of V, and

(ii) the range of T is a subspace of W.

EXAMPLE 3.58 Find the kernel and the range of the following linear transformations.

(i) $T: R^2 \to R^2$, $T(x, y) = (x - y, x + y)$

(ii) $T: R^3 \to R^2$, $T(x, y, z) = (x + y + z, x + y)$

(iii) $T: P_2 \to P_3$, $T(p(x)) = xp(x)$

Solution:

(i) Let $\mathbf{v} = [x, y] \in R^2$

Suppose $\mathbf{v} \in \ker(T)$

$$T(\mathbf{v}) = \mathbf{0}$$

$$T(x, y) = (0, 0)$$

$$(x - y, x + y) = (0, 0)$$
$$x - y = 0, \quad x + y = 0$$
$$x = 0, \quad y = 0$$
$$\therefore \qquad \mathbf{v} = [x, y] = [0, 0]$$

Hence, $\quad \ker(T) = \{[0, 0]\}$.

The range space of T is

$$\begin{aligned} R(T) &= \{T(\mathbf{v}) \mid \mathbf{v} = (x, y) \in R^2\} \\ &= \{(x - y, x + y) \mid x, y \in R\} \\ &= \{(x, x) + (-y, y) \mid x, y \in R\} \\ &= \{x(1, 1) + y(-1, 1) \mid x, y \in R\} \end{aligned}$$

(ii) Let $\mathbf{v} = [x, y, z]$ be a vector of R^3.
Suppose $\mathbf{v} \in \ker(T)$

$$T(\mathbf{v}) = \mathbf{0}$$
$$T(x, y, z) = (0, 0)$$
$$(x + y + z, x + y) = (0, 0)$$
$$x + y + z = 0, \qquad x + y = 0$$
$$y = -x, \quad z = 0$$

If $x = k$, then $y = -k$, $z = 0$

$$\mathbf{v} = [x, y, z] = [k, -k, 0]$$
$$\begin{aligned} \ker(T) &= \{\mathbf{v} \in R^3 \mid T(\mathbf{v}) = 0\} = \{(k, -k, 0) \mid k \in R\} \\ &= \{k(1, -1, 0) \mid k \in R\} \end{aligned}$$

The range space of T is

$$\begin{aligned} R(T) &= \{T(\mathbf{v}) \mid \mathbf{v} = [x, y, z] \in R^3\} \\ &= \{(x + y + z, x + y\} \mid x, y, z \in R\} \\ &= \{(x + y, x + y) + (z, 0) \mid x, y, z \in R\} \\ &= \{(x + y)(1, 1) + z(1, 0) \mid x, y, z \in R\} \\ &= \{t(1, 1) + z(1, 0) \mid t, z \in R\} \end{aligned}$$

(iii) Suppose $p(x) = p_0 + p_1x + p_2x^2 \in \ker(T)$

$$T(p(x)) = \mathbf{0}$$
$$x(p(x)) = \mathbf{0}$$
$$p_0x + p_1x^2 + p_2x^3 = 0.1 + 0\,x + 0\,x^2 + 0\,x^3$$
$$p_0 = 0, \quad p_1 = 0, \quad p_2 = 0$$
$$p(x) = \mathbf{0}$$

Therefore $\quad \ker(T) = \{p(x) \mid T(p(x)) = \mathbf{0}\} = \{\mathbf{0}\}$

The range space of T is

$$\begin{aligned} \mathrm{R}(T) &= \{T(p(x))\,|\,p(x) \in P_2\} \\ &= \{p_0x + p_1x^2 + p_2x^3\,|\,p_{0,}\ p_1, p_2 \in R\}. \end{aligned}$$

Recall that, every Euclidean linear transformation T: $R^n \to R^m$ can be expressed as

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x} \qquad \text{where } \mathbf{A} \text{ is an } m \times n \text{ matrix.}$$

The following theorem gives the information about the range space and the kernel of such linear transformations.

Theorem 3.9 [Null Space and Column Space of A]

If T: $R^n \to R^m$ is a linear transformation defined by the formula $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is an $m \times n$ matrix, then

(i) ker (T) = null $(\mathbf{A})$ (null space of $\mathbf{A}$)

(ii) R (T) = coll $(\mathbf{A})$ (column space of $\mathbf{A}$)

EXAMPLE 3.59 Find the kernel and the range of the following Euclidean linear transformations.

(i) T: $R^2 \to R^2$, $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ where $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$

(ii) T: $R^4 \to R^2$, $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A} = \begin{bmatrix} 1 & 4 & -2 & 2 \\ -1 & 2 & 2 & 1 \end{bmatrix}$

Solution:

(i) Let $\mathbf{v} = [x, y]$ be a vector of R^2.

Suppose $\mathbf{v} \in$ ker (T) (by Theorem 3.9)

$$\mathbf{A}\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 1 & 2 & : & 0 \\ 3 & -1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & : & 0 \\ 0 & -7 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & : & 0 \\ 0 & 1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & : & 0 \\ 0 & 1 & : & 0 \end{bmatrix}$$

Therefore the corresponding equations are

$$x = 0,\ y = 0$$

and $\quad [x, y] = [0, 0]$

$\therefore \quad$ ker $(T) = \{[0, 0]\}$

The range space of T is

$$\mathrm{R}(T) = \mathrm{col}\ (\mathbf{A}) \qquad \text{(by Theorem 3.9)}$$

$$= \left\{ x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} \,\middle|\, x_1, x_2 \in R \right\}$$

(ii) Let $\mathbf{v} = [x_1, x_2, x_3, x_4]$ be a vector of R^4.
Suppose $\mathbf{v} \in \ker(T)$

$\mathbf{v} \in \text{null}(\mathbf{A})$ (by Theorem 3.9)

$$\mathbf{Av} = \mathbf{0}$$

$$\begin{bmatrix} 1 & 4 & -2 & 2 \\ -1 & 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 1 & 4 & -2 & 2 & : & 0 \\ -1 & 2 & 2 & 1 & : & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 & : & 0 \\ 0 & 2 & 0 & 1 & : & 0 \end{bmatrix}$$

The linear equations for this system are

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 - 2x_3 = 0; \qquad 2x_2 + x_4 = 0$$

If $x_2 = s$, $x_3 = t$, then $x_1 = 2t$, $x_4 = -2s$. That is, $\mathbf{v} = [x_1, x_2, x_3, x_4] = [2t, s, t, -2s]$

$$\therefore \qquad \ker(T) = \{[2t, s, t, -2s] \mid s, t \in R)$$
$$= \{t(2, 0, 1, 0) + s(0, 1, 0, -2) \mid s, t \in R\}$$

The range of T is

$$R(T) = \text{col}(\mathbf{A}) \qquad \text{(by Theorem 3.9)}$$

$$= \left\{\alpha_1\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_2\begin{bmatrix} 4 \\ 2 \end{bmatrix} + \alpha_3\begin{bmatrix} -2 \\ 2 \end{bmatrix} + \alpha_4\begin{bmatrix} 2 \\ 1 \end{bmatrix} \;\middle|\; \alpha_i \in R\right\}$$

$$= \left\{\alpha_1\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2\alpha_2\begin{bmatrix} 2 \\ 1 \end{bmatrix} - 2\alpha_3\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_4\begin{bmatrix} 2 \\ 1 \end{bmatrix} \;\middle|\; \alpha_i \in R\right\}$$

$$= \left\{(\alpha_1 - 2\alpha_3)\begin{bmatrix} 1 \\ -1 \end{bmatrix} + (2\alpha_2 + \alpha_4)\begin{bmatrix} 2 \\ 1 \end{bmatrix} \;\middle|\; \alpha_i \in R\right\}$$

$$= \left\{\alpha\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta\begin{bmatrix} 2 \\ 1 \end{bmatrix} \;\middle|\; \alpha, \beta \in R\right\}.$$

In Theorem 3.8, we saw that the kernel and the range of T are vector spaces. Therefore, they have bases and hence dimensions. Thus, we have the following definitions.

Definition **Nullity and Rank of Linear Transformation**

If $T: V \to W$ is a linear transformation, then the dimension of kernel of T is called the *nullity* of T and the dimension of range of T is called the *rank* of T.

Notation: Nullity of T = nullity (T)

Rank of T = rank (T)

The following theorem is an extension of Theorem 3.9.

Theorem 3.9(a) [Rank and Nullity of Linear Transformation]

If $T: R^n \to R^m$ is a linear transformation defined by the formula, $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is an $m \times n$ matrix, then

(i) nullity T = nullity $\mathbf{A}$
(ii) rank T = rank $\mathbf{A}$

Theorem 3.10 [Rank-Nullity Theorem]

Let V be a finite dimensional vector space. If $T: V \to W$ be a linear transformation, then

$$\text{rank } [T] + \text{nullity } [T] = \dim V.$$

EXAMPLE 3.60 Find the rank and nullity of the following linear transformations. Hence verify the rank-nullity theorem.

(i) $T: R^2 \to R^2$, $T(x, y) = (x - y, x + y)$
(ii) $T: R^3 \to R^2$, $T(x, y, z) = (x + y + z, x + y)$
(iii) $T: P_2 \to P_3$, $T(p(x)) = xp(x)$
(iv) $T: R^2 \to R^2$, $T(x) = \mathbf{A}x$, where $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$
(v) $T: R^4 \to R^2$, $T(x) = \mathbf{A}x$ where $\mathbf{A} = \begin{bmatrix} 1 & 4 & -2 & 2 \\ -1 & 2 & 2 & 1 \end{bmatrix}$

Solution:

(i) See Example 3.58(i),

$$\ker(T) = \{[0, 0]\}$$
$$\text{nullity } (T) = \dim(\ker(T)) = 0$$
$$R(T) = \{x(1, 1) + y(-1, 1) \mid x, y \in R\}$$

Hence basis for R(T) is $B = \{[1, 1], [-1, 1]\}$

$\therefore$ $\quad$ rank $T = \dim(\text{R}(T)) = 2$

Now rank T + nullity $T = 2 + 0 = 2 = \dim V$ $\quad$ since $\dim V = \dim R^2 = 2$.

Hence the rank-nullity theorem stands verified.

(ii) See Example 3.58(iii)

$$\ker(T) = \{s(1, -1, 0) \mid s \in R\}$$

Hence the basis for $\quad \ker(T) = \{[1, -1, 0]\}$

$$\text{nullity } T = \dim(\ker(T)) = 1$$

The range space of T is

$$R(T) = \{t(1, 1) + z(0, 1) \mid t, z \in R\}$$

Hence the basis for $\quad R(T) = \{[1, 1], [0, 1]\}$

$$\text{rank } T = \dim (R(T)) = 2$$

$\therefore \quad$ rank T + nullity $T = 2 + 1 = 3 = \dim V \qquad$ since $\dim V = \dim R^3 = 3$

(iii) Look at Example 3.58(iii)

$$\ker (T) = \{\mathbf{0}\}; \quad \text{nullity } T = 0$$

and $\quad R(T) = \{p_0 + p_1x + p_2x^2 \mid p_0, p_1, p_2 \in R\}$

Hence $\quad$ basis for $R(T) = \{x, x, x^3\}$

and $\quad$ rank $T = \dim (R(T)) = 3$

$\therefore \quad$ rank T + nullity $T = 3 + 0 = 3 = \dim V \quad$ since $\dim V = \dim P_2 = 3$

(iv) See Example 3.59(i)

$$\ker (T) = \{[0, 0]\}$$

$$\text{nullity } T = 0$$

and $\quad R(T) = \left\{\alpha_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} \middle| \, \alpha_1, \alpha_2, \in R\right\}$

Hence $\quad$ basis for $R(T) = \left\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}$

and $\quad$ rank $T = \dim [R(T)] = 2$

$\therefore \quad$ rank T + nullity $T = 2 + 0 = 2 = \dim V \qquad$ since $\dim V = \dim R^2 = 2$

(v) Look at Example 3.59(ii)

$$\ker (T) = \{t\,(2, 0, 1, 0) + s\,(0, 1, 0, -2) \mid s, t \in R\}$$

Hence $\quad$ the basis for $\ker (T) = \{[2, 0, 1, 0], [0, 1, 0, -2]\}$

Hence $\quad$ nullity $T = \dim (\ker (T)) = 2$

Since $\quad R(T) = \{\alpha(1, -1) + \beta(2, 1) \mid \alpha, \beta \in R\}$

$\therefore \quad$ basis for $R(T) = \{[1, -1], [2, 1]\}$.

Thus $\quad$ rank $[T] = 2$

$\therefore \quad$ rank T + nullity $T = 2 + 2 = 4 = \dim V \qquad$ since $\dim V = \dim R^4 = 4$.

Hence the rank-nullity theorem is verified.

Theorem 3.11 [Properties of Linear Transformation]

Let $T: V \to W$ be a linear transformation, then the following statements are equivalent.

(i) T is one-to-one.
(ii) $\ker (T) = \{\mathbf{0}\}$.
(iii) nullity $T = 0$.

EXAMPLE 3.61 Find the rank and nullity of the linear transformation $T: R^2 \to R^3$,

$$T(x, y) = (x - y, 2y, x + 3y).$$

Solution: The linear transformation $T(x, y) = (x - y, 2y, x + 3y)$ is a one-to-one on R^2 (see Example 3.55(ii)).

By Theorem 3.11, ker $(T) = \{\mathbf{0}\}$. Therefore, nullity $T = 0$.

By the rank-nullity theorem,

$$\begin{aligned} \text{rank } T &= \dim V - \text{nullity } T \\ &= \dim R^2 - \text{nullity } T \\ &= 2 - 0 = 2 \end{aligned}$$

EXERCISE SET 5

1. Find the kernel and the range of the following linear transformations.
   (i) $T : R^2 \to R^2$, $T(x, y) = (y, x)$.
   (ii) $T : R^2 \to R^2$, $T(x, y) = (4x + 2y, 0)$.
   (iii) $T : R^2 \to R^3$, $T(x, y) = (x, x + y, x - y)$
   (iv) $T : R^3 \to R^2$, $T(x, y) = (x + y + z, x - y - z)$.

2. Find the kernel and the range of the following Euclidean linear transformations generated by a matrix.

$$T: R^2 \to R^2, \qquad T(\mathbf{x}) = \mathbf{Ax} \quad \text{where } \mathbf{A} = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

3. Let $T: R^2 \to R^2$ be the orthogonal projection on the line $y = x$.
   (i) Find the kernel of T.
   (ii) Is T one-to-one? Justify your conclusion.

4. Find the rank and the nullity of

$$T: M_{22} \to P_3, \qquad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a + b + 2d) + 0x + cx^2 + cx^3.$$

Hence verify the rank-nullity theorem.

5. Verify the rank-nullity theorem for the linear transformation

$$T: P_3 \to P_3, \qquad T(p(x)) = p'(x).$$

6. Find the nullity of the following linear transformations:
   (i) $T: R^4 \to R^6$ of rank four
   (ii) $T: P_3 \to P_3$ of rank one
   (iii) $T: R^6 \to R^3$ onto map
   (iv) $T: M_{22} \to P_3$ onto map

7. Answer the following questions:
   (i) If $T: R^n \to R^n$ is a one-to-one map, then what is the nullity of T?
   (ii) Is $T: R^n \to R^n$ one-to-one map if $\mathrm{R}(T) = n$?
   (iii) Is nullity $T = 0$ if $T: R^n \to R^m$, $m < n$?
   (iv) Is $T: R^n \to R^n$ one-to-one map if rank $T = n - 1$?

3.7 CHANGE OF BASIS AND SIMILARITY

Let V be a vector space with basis B. Suppose B' is another basis for V. Here we want to reconstruct vector space V with respect to new basis of B', i.e. we want to change the basis of V. In other words, we have to find new coordinates of each vector $\mathbf{v}$ of V with respect to new basis B'. This problem can be simplified by finding the relationship between $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$.

To solve this problem, consider the identity operator on V that maps each vector of V onto itself, that is, $I: V \to V$, $I(\mathbf{v}) = \mathbf{v}$ and find the matrix $[I]_{B,B'}$.

Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_n'\}$ be bases for vector space V. By the definition of associated matrix with the linear transformation,

$$[I(\mathbf{v})]_B = [I]_{B,B'}\,[\mathbf{v}]_{B'}$$

where $[I]_{B,B'} = \left[[I(\mathbf{u}_1')]_B\,[I(\mathbf{u}_2')]_B \cdots [I(\mathbf{u}_n')]_B\right] = \left[[(\mathbf{u}_1')]_B\,[(\mathbf{u}_2')]_B \cdots [(\mathbf{u}_n')]_B\right]$, since I is the identity operator.

$$[I(\mathbf{v})]_B = [[\mathbf{u}_1']_B\,[\mathbf{u}_2']_B \cdots [\mathbf{u}_n']_B]\,[\mathbf{v}]_{B'}$$

$$[\mathbf{v}]_B = [[\mathbf{u}_1']_B\,[\mathbf{u}_2']_B \cdots [\mathbf{u}_n']_B]\,[\mathbf{v}]_{B'} \qquad \text{since } I(\mathbf{v}) = \mathbf{v}$$

Definition: Transition Matrix

If B and B' are the two bases for vector space V and $\mathbf{P}$ is a matrix that relates $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ by the equation. $[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$ for every $\mathbf{v} \in V$, then $\mathbf{P}$ is called a transition matrix from B' to B. As per the above discussion, the transition matrix $\mathbf{P}$ can be expressed as

$$\mathbf{P} = [I]_{B,B'} = [[\mathbf{u}_1']_B\,[\mathbf{u}_2']_B \cdots [\mathbf{u}_n']_B]$$

where $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_n'\}$.

EXAMPLE 3.62 Find the transition matrix from B' to B.

If $B = \{[1, 0], [0, 1]\}$, $B' = \{[1, -2], [3, 8]\}$ for R^2. Compute $[\mathbf{v}]_B$ when $[\mathbf{v}]_{B'} = [2, 3]$

Solution: Say $\mathbf{u}_1 = [1, 0]$, $\mathbf{u}_2 = [0, 1]$ and $\mathbf{u}_1' = [1, -2]$, $\mathbf{u}_2' = [3, 8]$

By the definition of transition matrix, $\mathbf{P}$ is given by the formula

$$\mathbf{P} = [[\mathbf{u}_1']_B\,[\mathbf{u}_2']_B]$$

Since $\mathbf{u}_1' = [1, -2]$, $\mathbf{u}_2' = [3, 8]$,

$$\mathbf{u}_1' = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 = 1(1, 0) + (-2)(0, 1); \qquad \mathbf{u}_2' = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 = 3(1, 0) + 8(0, 1)$$

$$[u_1']_B = \begin{bmatrix} 1 \\ -2 \end{bmatrix}; \qquad [u_2']_B = \begin{bmatrix} 3 \\ 8 \end{bmatrix};$$

Thus the transition matrix

$$\mathbf{P} = [[u_1']_B\,[u_2']_B] = \begin{bmatrix} 1 & 3 \\ -2 & 8 \end{bmatrix}.$$

By the property of transition matrix,

$$[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$$

$$\therefore \qquad [\mathbf{v}]_B = \begin{bmatrix} 1 & 3 \\ -2 & 8 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}_{B'} = \begin{bmatrix} 11 \\ 20 \end{bmatrix}$$

EXAMPLE 3.63 Find the transition matrix from B to B', if $B = \{[1, 1], [0, 2]\}$ and $B' = \{[2, 1], [-3, 4]\}$ for R^2. Compute $[\mathbf{v}]_B$ when $[\mathbf{v}]_{B'} = [0, 1]$.

Solution: Say $\mathbf{u}_1 = [1, 1]$, $\mathbf{u}_2 = [0, 2]$ and $\mathbf{u}_1' = [2, 1]$, $\mathbf{u}_2' = [-3, 4]$

By the definition of transition matrix, **P** is given by the formula

$$\mathbf{P} = [[\mathbf{u}_1']_B\ [[\mathbf{u}_2']_B]$$

Since $\qquad \mathbf{u}_1' = [2, 1]$, $\mathbf{u}_2' = [-3, 4]$,

$$\mathbf{u}_1' = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2; \qquad \mathbf{u}_2' = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2$$

or

$$(2, 1) = \alpha_1(1, 1) + \alpha_2(0, 2); \qquad (-3, 4) = \beta_1(1, 1) + \beta_2(0, 2)$$

$$\mathbf{u}_1' = 2\mathbf{u}_1 + \left(-\frac{1}{2}\right)\mathbf{u}_2; \qquad \mathbf{u}_2' = (-3)\mathbf{u}_1 + \left(\frac{7}{2}\right)\mathbf{u}_2$$

$$[\mathbf{u}_1']_B = \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix}; \qquad [\mathbf{u}_2']_B = \begin{bmatrix} -3 \\ \frac{7}{2} \end{bmatrix}$$

The transition matrix $\mathbf{P} = [[\mathbf{u}_1']_B\ [\mathbf{u}_2']_B] = \begin{bmatrix} 2 & -3 \\ -\frac{1}{2} & \frac{7}{2} \end{bmatrix}$

By the transition matrix,

$$[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$$

$$= \begin{bmatrix} 2 & -3 \\ -\frac{1}{2} & \frac{7}{2} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ \frac{7}{2} \end{bmatrix}$$

EXAMPLE 3.64 Find the transition matrix from B' to B, if $B = \{[1, 0, 0], [1, 1, 0], [1, 1, 1]\}$ and $B' = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ for R^3. Compute $[\mathbf{v}]_B$ when $[\mathbf{v}]_{B'} = [1, 2, 3]$.

Solution: Say $\mathbf{u}_1' = [1, 0, 0]$, $\mathbf{u}_2' = [0, 1, 0]$, $\mathbf{u}_3' = [0, 0, 1]$.

The transition matrix **P** is given by the formula

$$\mathbf{P} = [[\mathbf{u}_1']_B\ [\mathbf{u}_2']_B\ [\mathbf{u}_3']_B]$$

Since $\mathbf{u}_1' = [1, 0, 0]$, $\mathbf{u}_2' = [0, 1, 0]$, $\mathbf{u}_3' = [0, 0, 1]$,

$$\mathbf{u}_1' = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3$$

$$\mathbf{u}_2' = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \beta_3\mathbf{u}_3$$

$$\mathbf{u}_3' = \gamma_1\mathbf{u}_1 + \gamma_2\mathbf{u}_2 + \gamma_3\mathbf{u}_3$$

$$(1, 0, 0) = \alpha_1(1, 0, 0) + \alpha_2(1, 1, 0) + \alpha_3(1, 1, 1)$$
$$(0, 1, 0) = \beta_1(1, 0, 0) + \beta_2(1, 1, 0) + \beta_3(1, 1, 1)$$
$$(0, 0, 1) = \gamma_1(1, 0, 0) + \gamma_2(1, 1, 0) + \gamma_3(1, 1, 1)$$

By solving the above system of linear equations, we get

$$\mathbf{u}_1' = 1\mathbf{u}_1 + 0\mathbf{u}_2 + 0\mathbf{u}_3; \qquad \mathbf{u}_2' = (-1)\mathbf{u}_1 + 1\mathbf{u}_2 + 0\mathbf{u}_3; \qquad \mathbf{u}_3' = 0\mathbf{u}_1 + (-1)\mathbf{u}_2 + 1\mathbf{u}_3$$

Hence $$[\mathbf{u}_1']_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad [\mathbf{u}_2']_B = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad [\mathbf{u}_3']_B = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

The transition matrix $\mathbf{P} = [[\mathbf{u}_1']_B\ [\mathbf{u}_2']_B\ [\mathbf{u}_3']_B]$

$$= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

For $[\mathbf{v}]_{B'} = [1, 2, 3] = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$$

$$= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$$

EXAMPLE 3.65 Find the transition matrix from B' to B, if bases $B = \{1 + 2x, 2 + x\}$ and $B' = \{1, x\}$ for P_1. Compute $[\mathbf{v}]_B$ when $[\mathbf{v}]_{B'} = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$.

Solution: Let $\mathbf{u}_1 = 1 + 2x$, $\mathbf{u}_2 = 2 + x$ and $\mathbf{u}_1' = 1$, $\mathbf{u}_2' = x$

The transition matrix $\mathbf{P}$ is given by the formula, $\mathbf{P} = [[\mathbf{u}_1']_B\ [\mathbf{u}_2']_B]$

Since $\mathbf{u}_1' = 1, \mathbf{u}_2' = x$

$$\mathbf{u}_1' = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2; \qquad \mathbf{u}_2' = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2$$

or $$1 = \alpha_1(1 + 2x) + \alpha_2(2 + x); \qquad x = \beta_1(1 + 2x) + \beta_2(2 + x)$$

By solving the above system of linear equations,

$$\mathbf{u}_1' = \left(-\frac{1}{3}\right)\mathbf{u}_1 + \left(\frac{2}{3}\right)\mathbf{u}_2; \qquad \mathbf{u}_2' = \left(\frac{2}{3}\right)\mathbf{u}_1 + \left(-\frac{1}{3}\right)\mathbf{u}_2$$

Hence

$$[\mathbf{u}_1']_B = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}; \qquad [\mathbf{u}_2']_B = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}$$

The transition matrix, $\mathbf{P} = [[\mathbf{u}_1']_B\ [\mathbf{u}_2']_B]$.

$$= \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

For $\quad [\mathbf{v}]_{B'} = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$

$$[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$$

$$[\mathbf{v}]_B = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} -3 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix} = 5 - 4x$$

Theorem 3.12 [Transition Matrix]

If V is a finite dimensional vector space with bases B and B' and $\mathbf{P}$ is the transition matrix from B' to B, then $\mathbf{P}$ is invertible and $\mathbf{P}^{-1}$ is the transition matrix from B to B', that is,

$$[\mathbf{v}]_{B'} = \mathbf{P}^{-1}\,[\mathbf{v}]_B$$

EXAMPLE 3.66 Find the transition matrix from B to B',

if $B = \{[1, 0], [0, 1]\}$, $B' = \{[1, -2], [3, 8]\}$ for R^2. Compute $[\mathbf{v}]_{B'}$ when $[\mathbf{v}]_B = [11, 20]$

Solution: In Example 3.62 we calculated the transition matrix $\mathbf{P}$ from B' to B, i.e.

$$\mathbf{P} = \begin{bmatrix} 1 & 3 \\ -2 & 8 \end{bmatrix}$$

Therefore, by Theorem 3.12 the transition matrix from B to B' is

$$\mathbf{P}^{-1} = \frac{1}{|\mathbf{P}|}\,\text{adj}\,\mathbf{P} = \frac{1}{14}\begin{bmatrix} 8 & -3 \\ 2 & 1 \end{bmatrix}$$

For $\quad [\mathbf{v}]_B = [11, 20]$

$$[\mathbf{v}]_{B'} = \mathbf{P}^{-1}\,[\mathbf{v}]_B$$

$$= \frac{1}{14}\begin{bmatrix} 8 & -3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 11 \\ 20 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

EXAMPLE 3.67 Find the transition matrix from B to B',
if bases $B = \{1 + 2x, 2 + x\}$ and $B' = \{1, x\}$ for P_1. Compute $[\mathbf{v}]_{B'}$ when $[\mathbf{v}]_B = 5 - 4x$.

Solution: In Example 3.65, the transition matrix $\mathbf{P}$ from B' to B is $\mathbf{P} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$.

Therefore by Theorem 3.12, the transition matrix from B to B' is

$$\mathbf{P}^{-1} = \frac{1}{|\mathbf{P}|} \text{adj } \mathbf{P}$$

$$= (-3)\begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

For $\quad [\mathbf{v}]_B = 5 - 4x = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$

$$[\mathbf{v}]_{B'} = \mathbf{P}^{-1} [\mathbf{v}]_B$$

$$= \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

$$= -3 + 6x$$

Similarity

In the previous section, we calculated the relationship between two bases B and B' of the vector space V under the identity linear operator on V. We got the following equation $[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$ where $\mathbf{P} = [I]_{B,B'}$ is the transition matrix from B' to B. Here we will try to expand this concept for the general linear operator T on V.

In other words, if $T: V \to V$ is a linear operator on V, then what is the relationship between $[T]_B$ and $[T]_{B'}$. After finding this relationship, we will see how this concept works on the square matrices of the same order.

To find such a relationship, we consider the following three linear operators:

Identity operator $\quad I: V_{B'} \to V_B, \quad [\mathbf{u}]_B = [I]_{B,B'}[\mathbf{u}]_{B'}$

Linear operator $\quad T: V_B \to V_B, \quad [T(\mathbf{u})]_B = [T]_B [\mathbf{u}]_B$

Identity operator $\quad I': V_B \to V_{B'}, \quad [\mathbf{u}]_{B'} = [I]_{B',B} [\mathbf{u}]_B$

where V_B and $V_{B'}$ represent the vector space V with bases B and B' respectively.

By taking the composition of these operators, we get the linear operators $I' \text{o} T \text{o} I: V_{B'} \to V_{B'}$

$$I' \text{o} T \text{o} I\, [\mathbf{u}]_{B'} = [I' \text{o} T]\, [I(\mathbf{u})_{B'}]$$

$$= [I' \text{o} T]\, ([I]_{B,B'}[\mathbf{u}]_{B'})$$

$$= [I' \text{o} T][\mathbf{u}]_B$$

Since $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$,

$$T(\mathbf{u}_1) = T(1, 0) = (1, 1); \qquad T(\mathbf{u}_2) = T(1, 0) = (1, -1)$$

$$[T(\mathbf{u}_1)]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \qquad [T(\mathbf{u}_2)]_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix};$$

$$\begin{aligned} [T]_B &= [[T(\mathbf{u}_1)]_B\ [T(\mathbf{u}_2)]_B] \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

(iii) Now we will try to find the matrix $[T]_{B'}$ of linear transformation T with respect to basis B'. By the definition,

$$[T]_{B'} = [[T(\mathbf{u}_1')]_{B'}\ [T(\mathbf{u}_2')]_{B'}]$$

Since $T(x_{1,} x_2) = (x_1 - x_2, x_1 + x_2)$,

$$T(\mathbf{u}_1') = T(1, 1) = (0, 2); \qquad T(\mathbf{u}_2') = T(1, -1) = (2, 0)$$

$$T(\mathbf{u}_1') = \alpha_1\mathbf{u}_1' + \alpha_2\mathbf{u}_2' = 1\mathbf{u}_1' + (-1)\mathbf{u}_2'; \qquad T(\mathbf{u}_2') = \beta_1\mathbf{u}_1' + \beta_2\mathbf{u}_2' = 1\mathbf{u}_1' + 1\mathbf{u}_2'$$

Hence $[T(\mathbf{u}_1')]_{B'} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}; \qquad [T(\mathbf{u}_2')]_{B'} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$[T]_{B'} = [(T(\mathbf{u}_1'))_{B'}\ (T(\mathbf{u}_2'))_{B'}]$$

Hence $$[T]_{B'} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{i}$$

The right-hand side of the result of Theorem 3.13 is

$$\begin{aligned} \mathbf{P}^{-1}[T]_B\,\mathbf{P} &= -1/2\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= -1/2\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \\ &= -1/2\begin{bmatrix} -2 & -2 \\ 2 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= [T]_{B'} \end{aligned}$$

Hence, we have verified the result of Theroem 3.13.

Similar Matrices

Two square matrices **A** and **B** of order n are called *similar matrices* if there is an invertible matrix **C** such that $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$.

The following theorem is important to find such a matrix **C**.

Theorem 3.14 [Similar Matrices]

Two matrices **A** and **B** are similar if and only if they represent the same linear transformation.

EXAMPLE 3.69

Prove that $\begin{bmatrix} 10 & 2 \\ 2 & 7 \end{bmatrix}$ is similar to $\begin{bmatrix} 11 & 0 \\ 0 & 6 \end{bmatrix}$ via the non-singular matrix $\begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}$.

Solution: Let $\mathbf{A} = \begin{bmatrix} 11 & 0 \\ 0 & 6 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 10 & 2 \\ 2 & 7 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}$.

We want to show that $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$.

RHS:

$$\begin{aligned} \mathbf{C}^{-1}\mathbf{AC} &= (-1)\begin{bmatrix} -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} 11 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} \\ &= (-1)\begin{bmatrix} -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \frac{22}{\sqrt{5}} & \frac{11}{\sqrt{5}} \\ \frac{6}{\sqrt{5}} & -\frac{12}{\sqrt{5}} \end{bmatrix} \\ &= (-1)\begin{bmatrix} -10 & -2 \\ -2 & -7 \end{bmatrix} \\ &= \begin{bmatrix} 10 & 2 \\ 2 & 7 \end{bmatrix} = \mathbf{B} \ \ (i.e. \text{ LHS}) \end{aligned}$$

Definition Invariant Under Similarity

A property of square matrices is said to be a *similarity invariant* or *invariant under similarity* if it is shared by any two similar matrices.

EXAMPLE 3.70 Prove that similar matrices have the same determinant or the determinant of similar matrices is similarity invariant.

Solution: If the matrix **B** is similar to matrix **A** via a non-singular matrice **C**, then $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$. We want to show that **A** and **B** have the same determinant.

Since $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$.

$$\begin{aligned} \det \mathbf{B} &= \det (\mathbf{C}^{-1}\mathbf{AC}) \\ &= \det \mathbf{C}^{-1} \det \mathbf{A} \det \mathbf{C} \end{aligned}$$

5. Prove that $\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$ is similar to $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ via the non-singular matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

6. Show that the matrices are $\begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$ are similar.

SUMMARY

Linear Transformation Let V and W be two vector spaces with the same sets of scalars. Then a function $T: V \to W$ is called a *linear transformation* of V into W if it satisfies the following properties:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, for all $\mathbf{u}$ and $\mathbf{v}$ in V

(ii) $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$, for every $\mathbf{v}$ in V and every scalar α

Definition: Euclidean Vector Spaces

The function $T: R^n \to R^m$ is called linear transformation if

$$T(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_m)$$

where

$$\begin{aligned} u_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ u_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots &\quad \vdots \qquad + \cdots + \quad \vdots \\ u_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \qquad 1 \le i \le m,\ 1 \le j \le n \end{aligned}$$

Linear Operator If $V = W$, then the linear transformation $T: V \to V$ is called a *linear operator* on V.

Theorem Let $T : V \to W$ be a linear transformation. If $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n$ are vectors in V and $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ are scalars, then

$$T(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 + \cdots + \alpha_n\mathbf{x}_n) = \alpha_1 T(\mathbf{x}_1) + \alpha_2 T(\mathbf{x}_2) + \alpha_3 T(\mathbf{x}_3) + \cdots + \alpha_n T(\mathbf{x}_n).$$

Corollary Let $T: V \to W$ be a linear transformation and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ be a basis for V.

If
$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \cdots + \alpha_n\mathbf{v}_n$$
is a vector of V, then
$$T(\mathbf{v}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \alpha_3 T(\mathbf{v}_3) + \cdots + \alpha_n T(\mathbf{v}_n).$$

Corollary [Properties of Linear Transformation]

If $T: V \to W$ is a linear transformation, then

(i) $T(\mathbf{0}) = \mathbf{0}$ (ii) $T(-\mathbf{u}) = -T(\mathbf{u})$

(iii) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all $\mathbf{u}$ and $\mathbf{v}$ in V.

Algorithm to Find the Matrix Associated to Linear Transformation Let V and W be vector spaces. Suppose the sets $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ are bases for the vector spaces V and W respectively and T is linear transformation from V to W. Then the matrix associated with the linear transformation can be calculated using the following algorithm.

Step 1 Find the images of basis vectors of B under T, that is, find $T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$

Step 2 Find the co-ordinates of images with respect to basis B', that is, express the images as a linear combinations of vectors of basis B'. In other words,

$$T(\mathbf{e}_k) = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

$$[T(\mathbf{e}_k)]_{B'} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

Find the αs for each k, $1 \le k \le n$.

Step 3 Put them as columns of an $m \times n$ matrix

$$\mathbf{A} = [[T(\mathbf{e}_1)]_{B'} \ \ [T(\mathbf{e}_2)]_{B'} \ldots [T(\mathbf{e}_n)]_{B'}].$$

Addition and Scalar Multiplication of Linear Transformation Let S: $V \to W$ and T: $V \to W$ be linear transformations. If α is any scalar, then the addition $S + T$ and scalar multiplication (αT) are defined by the formulae;

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}), \qquad \text{for } \mathbf{v} \in V$$

$$(\alpha T)(\mathbf{v}) = \alpha T(\mathbf{v}), \qquad \text{for all } \mathbf{v} \in V \text{ and any scalar } \alpha.$$

Theorem The set $L(V, W)$ of all linear transformations from V into W is a vector space with operations,

$$\textit{Addition: } (T + S)(\mathbf{v}) = T(\mathbf{v}) + S(\mathbf{v}), \qquad \text{for } S, T \,\varepsilon\, L(V, W)$$

$$\textit{Scalar multiplication: } (\alpha T)(\mathbf{v}) = \alpha T(\mathbf{v}), \qquad \text{for all scalars } \alpha.$$

Composition Operation Let V, U, W be vector spaces. If T_1: $V \to U$ and T_2: $U \to W$ are linear transformations, then composition of T_2 with T_1 is the function $T_2 o T_1$: $V \to W$ defined by the equation

$$(T_2 o T_1)(\mathbf{v}) = T_2(T_1(\mathbf{v})), \qquad \text{for all } \mathbf{v} \text{ in } V.$$

Theorem If T_1: $V \to U$ and T_2: $U \to W$ are linear transformations, then $T_2 o T_1$: $V \to W$ is also a linear transformation.

Theorem

(i) Let T_1: $V \to U$ and T_2: $U \to W$ be two linear transformations. If B, B'' and B' are bases for V, U and W respectively, then $[T_2 o T_1]_{B', \mathrm{B}} = [T_2]_{B', B''}[T_1]_{B'', B}$.

(ii) Let T_1: $V \to U$ and T_2: $U \to W$ and T_3: $W \to X$ be linear transformation. If B, B'' , B''' and B' are bases for the vector space V, U, W and X respectively, then

$$[T_3 o T_2 o T_1]_{B', B} = [T_3]_{B', B'''}[T_2]_{B''', B''}[T_1]_{B'', B}.$$

One-to-One Linear Transformation If the linear transformation T: $V \to W$ maps distinct vectors in V into distinct vectors in W, then T is called *one-to-one linear transformation.*

Theorem If T: $V \to W$ be a linear transformation and dim $V = n$, then the following statements are equivalent.

(i) T is one-to-one on V.

(ii) If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are independent elements in V, then $T(\mathbf{v}_1)$,$T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ are independent elements in $T(V)$.

(iii) If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for V, then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)\}$ is a basis for $T(V)$.

(iv) dim $T(V) = n$.

Inverse of Linear Transformation Let T: $V \to W$ be a one-to-one linear transformation, then the inverse of T is a map T^{-1}: $T(V) \to V$ such that $T^{-1}(\mathbf{w}) = \mathbf{v}$ if $T(\mathbf{v}) = \mathbf{w}$ for all $\mathbf{v}$ in V and $\mathbf{w}$ in $T(V)$.

Change of Basis If B and B' are the two bases for vector space V and $\mathbf{P}$ is a matrix that relates $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ by the equation $[\mathbf{v}]_B = \mathbf{P}[\mathbf{v}]_{B'}$ for every $\mathbf{v} \in V$, then $\mathbf{P}$ is called a *transition matrix* from B' to B.

Moreover the transition matrix $\mathbf{P}$ can be expressed as

$$\mathbf{P} = [I]_{B,B'} = [[u_1']_B \; [u_2']_B \cdots [u_n']_B]$$

where $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_n'\}$

Theorem [Transition Matrix]

If V is a finite dimensional vector space with bases B and B' and $\mathbf{P}$ is the transition matrix from B' to B, then $\mathbf{P}$ is invertible and $\mathbf{P}^{-1}$ is the transition matrix from B to B', that is,

$$[v]_{B'} = \mathbf{P}^{-1}[v]_B$$

Theorem If B and B' are the bases for the vector space V and $T: V \to V$ is a linear operator on V, then $[T]_{B'} = \mathbf{P}^{-1}[T]_B\,\mathbf{P}$ where $\mathbf{P}$ is the transition matrix from B' to B.

Similar Matrices Two square matrices $\mathbf{A}$ and $\mathbf{B}$ of order n are called *similar matrices* if there is an invertible matrix $\mathbf{C}$ such that $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$.

Theorem [Similar Matrices]

Two matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if and only if they represent the same linear transformation.

Similarity Invariant A property of square matrices is said to be a similarity invariant or invariant under similarity if it is shared by any two similar matrices.

4 Inner Product Spaces

4.1 INNER PRODUCT SPACE: DEFINITIONS AND CONCEPTS

In the Euclidean space R^n, we studied the concept of dot product which is used to find the length of a vector, the angle between vectors and orthogonality of vectors. Now, in this chapter, we will generalize the concept of dot product for the general vector space which will be called *inner product*. Also, we will define length, angle and orthogonality in terms of inner product, which will help us to understand the geometry of the space with respect to inner product.

Inner Product

In this section, we will give the precise definition of inner product on a general real vector space V. Also, we will see some standard properties of inner product.

Definition: Inner Product

A function defined from $V \times V$ into R is called inner product on a real vector space V if it assigns a unique real number written $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V that satisfies the following axioms for all vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all real scalars α.

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (Commutative/symmetry)

(ii) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (Distributive/linearity)

(iii) $\langle \alpha\mathbf{u}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$ (Homogeneity)

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ (Positivity) and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Definition: Inner Product Space

A real vector space with an inner product is called an inner product space.

EXAMPLE 4.1 Consider the real vector space R^n. If the function $\langle \mathbf{u}, \mathbf{v} \rangle$ (from $R^n \times R^n$ into R) is defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 + \cdots + u_n v_n$$

where

$$\mathbf{u} = [u_1, u_2, \ldots, u_n] \quad \text{and} \quad \mathbf{v} = [v_1, v_2, \ldots, v_n],$$

then show that $\langle \mathbf{u}, \mathbf{v} \rangle$ is inner product on R^n.

Solution: Let us verify that $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3 + \cdots + u_n v_n$ satisfies all the axioms of inner product.

(i) Let $\mathbf{u} = [u_1, u_2, \ldots, u_n]$, $\mathbf{v} = [v_1, v_2, \ldots, v_n] \in R^n$

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= u_1 v_1 + u_2 v_2 + u_3 v_3 + \cdots + u_n v_n \\ &= v_1 u_1 + v_2 u_2 + \cdots + v_n u_n \\ &= \langle \mathbf{v}, \mathbf{u} \rangle \end{aligned}$$

Hence $$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

(ii) Let $\mathbf{w} = [w_1, w_2, \ldots, w_n] \in R^n$

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = (u_1 + v_1) w_1 + (u_2 + v_2) w_2 + \cdots + (u_n + v_n) w_n$$

since $u + v = (u_1 + v_1, \ldots, u_n + v_n)$

$$\begin{aligned} &= u_1 w_1 + v_1 w_1 + u_2 w_2 + v_2 w_2 + \cdots + u_n w_n + v_n w_n \\ &= (u_1 w_1 + u_2 w_2 + \cdots + u_n w_n) + (v_1 w_1 + v_2 w_2 + \cdots + v_n w_n) \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

Hence $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

(iii) For any real scalar α,

$$\langle \alpha \mathbf{u}, \mathbf{v} \rangle = (\alpha u_1) v_1 + (\alpha u_2) v_2 + \cdots + (\alpha u_n) v_n \qquad \text{since } \alpha \mathbf{u} = [\alpha u_1, \alpha u_2, \ldots, \alpha u_n] \in R^n$$

$$\begin{aligned} &= \alpha (u_1 v_1 + u_2 v_2 + \cdots + u_n v_n) \\ &= \alpha \langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

Hence $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$

If $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, then $u_1^2 + u_2^2 + \cdots + u_n^2 = 0$

i.e. $$u_1 = 0, u_2 = 0, \ldots, u_n = 0$$

or $$\mathbf{u} = [u_1, u_2, \ldots, u_n] = [0, 0, \ldots, 0], \text{ i.e. } \mathbf{u} = \mathbf{0}$$

Therefore the given function $\langle \mathbf{u}, \mathbf{v} \rangle$ is an inner product in R^n; in other words R^n is an inner product space.

Note: The above inner product is known as *Euclidean inner product* on R^n and R^n is called *Euclidean* n*-space*.

EXAMPLE 4.2 Show that the function from $R^2 \times R^2$ into R is defined by

$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_1v_2 + u_2v_1 + 3u_2v_2$ where $\mathbf{u} = [u_1, u_2]$ and $\mathbf{v} = [v_1, v_2]$ is an inner product on R^2.

Solution: We will show that $\langle \mathbf{u}, \mathbf{v} \rangle$ satisfies all the axioms of inner product.

(i) Let $\mathbf{u} = [u_1, u_2]$, $\mathbf{v} = [v_1, v_2] \in R^2$

$$\begin{aligned}\langle \mathbf{v}, \mathbf{u} \rangle &= v_1u_1 + v_1u_2 + v_2u_1 + 3v_2u_2 \\ &= u_1v_1 + u_2v_1 + u_1v_2 + 3u_2v_2 \\ &= \langle \mathbf{u}, \mathbf{v} \rangle\end{aligned}$$

That is, $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(ii) Let $\mathbf{w} = [w_1, w_2] \in R^2$

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = (u_1 + v_1)w_1 + (u_1 + v_1)w_2 + (u_2 + v_2)w_1 + 3(u_2 + v_2)w_2$$

Since $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2)$

$$\begin{aligned}&= u_1w_1 + v_1w_1 + u_1w_2 + v_1w_2 + u_2w_1 + u_2w_1 + 3u_2w_2 + 3v_2w_2 \\ &= (u_1w_1 + u_1w_2 + u_2w_1 + 3u_2w_2) + (v_1w_1 + v_1w_2 + v_2w_1 + 3v_2w_2) \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle\end{aligned}$$

That is $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

(iii) For any scalar α,

$$\begin{aligned}\langle \alpha\mathbf{u}, \mathbf{v} \rangle &= (\alpha u_1)v_1 + (\alpha u_1)v_2 + (\alpha u_2)v_1 + 3(\alpha u_2)v_2 \qquad \text{Since } \alpha\mathbf{u} = [\alpha u_1, \alpha u_2] \\ &= \alpha(u_1v_1 + u_1v_2 + u_2v_1 + 3u_2v_2) \\ &= \alpha\langle \mathbf{u}, \mathbf{v} \rangle\end{aligned}$$

That is, $\langle \alpha\mathbf{u}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$

(iv)

$$\begin{aligned}\langle \mathbf{u}, \mathbf{u} \rangle &= u_1^2 + 2u_1u_2 + 3u_2^2 \\ &= (u_1^2 + 2u_1u_2 + u_2^2) + 2u_2^2 \\ &= (u_1 + u_2)^2 + 2u_2^2 \geq 0\end{aligned}$$

That is, $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$.

If $\langle \mathbf{u}, \mathbf{u} \rangle = 0$

then $(u_1 + u_2)^2 + 2u_2^2 = 0$

i.e. $u_1 = 0$, $u_2 = 0$

then $\mathbf{u} = \mathbf{0}$

Therefore the given function $\langle \mathbf{u}, \mathbf{v} \rangle$ is an inner product on R^2 and R^2 is an inner product space.

i.e. $\qquad p_0 = 0, p_1 = 0, \ldots, p_n = 0$

or $\qquad p(x) = 0 + 0x + \cdots + 0x^n$

That is, $\qquad \langle \mathbf{p}, \mathbf{p} \rangle = 0$ if and only if $\mathbf{p} = \mathbf{0}$.

Hence $\langle \mathbf{p}, \mathbf{q} \rangle$ is an inner product on the vector space P_n and so P_n is an inner product space.

EXAMPLE 4.5 Consider the vector space $P_2[-1, 1]$ of all polynomials over R of degree at most 2. If the function $\langle \mathbf{p}, \mathbf{q} \rangle$ is defined by $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} p(x)\, q(x) dx$ where $p(x) = p_0 + p_1x + p_2x^2$ and $q(x) = q_0 + q_1x + q_2x^2$, then show that P_2 is an inner product space with inner product $\langle \mathbf{p}, \mathbf{q} \rangle$.

Solution: We will verify the four axioms of inner product for the given $\langle \mathbf{p}, \mathbf{q} \rangle$.

(i) By using the results of integration,

$$\begin{aligned} \langle \mathbf{q}, \mathbf{p} \rangle &= \int_{-1}^{1} q(x)\; p(x)\, dx \\ &= \int_{-1}^{1} p(x)\, q(x)\, dx \\ &= \langle \mathbf{p}, \mathbf{q} \rangle \end{aligned}$$

That is, $\qquad \langle \mathbf{p}, \mathbf{q} \rangle = \langle \mathbf{q}, \mathbf{p} \rangle$

(ii) Let $r(x) = r_0 + r_1x + r_2x^2 \in P_2$

$$\begin{aligned} \langle \mathbf{p} + \mathbf{q}, \mathbf{r} \rangle &= \int_{-1}^{1} (p+q)\,(x) r(x) dx \\ &= \int_{-1}^{1} (p(x) + q(x))\, r(x) dx \\ &= \int_{-1}^{1} (p(x)r(x) + q(x)r(x))\, dx \\ &= \int_{-1}^{1} p(x)r(x)dx + \int_{-1}^{1} q(x)r(x)dx \\ &= \langle \mathbf{p}, \mathbf{r} \rangle + \langle \mathbf{q}, \mathbf{r} \rangle \end{aligned}$$

That is, $\qquad \langle \mathbf{p} + \mathbf{q}, \mathbf{r} \rangle = \langle \mathbf{p}, \mathbf{r} \rangle + \langle \mathbf{q}, \mathbf{r} \rangle$

(iii) For any scalar α,

$$\langle \alpha\mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} (\alpha\, p)\,(x) q(x) dx$$

$$= w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n$$
$$= \langle \mathbf{u}, \mathbf{v} \rangle$$

Hence $\quad \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(ii) Let $z = [z_1, z_2, \ldots, z_n) \in R^n$

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{z} \rangle = w_1(u_1 + v_1)z_1 + w_2(u_2 + v_2)z_2 + \cdots + w_n(u_n + v_n)z_n$$

since $u + v = (u_1 + v_1) + (u_2 + v_2) + \cdots + (u_n + v_n)$

$$= w_1u_1z_1 + w_1v_1z_1 + w_2u_2z_2 + w_2v_2z_2 + \cdots + w_nu_nz_n + w_nv_nz_n$$
$$= (w_1u_1z_1 + w_2u_2z_2 + \cdots + w_nu_nz_n) + (w_1v_1z_1 + w_2v_2z_2 + \cdots + w_nv_nz_n)$$
$$= \langle \mathbf{u}, \mathbf{z} \rangle + \langle \mathbf{v}, \mathbf{z} \rangle$$

Hence $\langle \mathbf{u} + \mathbf{v}, \mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{z} \rangle + \langle \mathbf{v}, \mathbf{z} \rangle$

(iii) For any scalar α,

$$\langle \alpha\mathbf{u}, \mathbf{v} \rangle = w_1(\alpha u_1)v_1 + w_2(\alpha u_2)v_2 + \cdots + w_n(\alpha u_n)v_n$$

since $\alpha u = (\alpha u_1, \alpha u_2, \ldots, \alpha u_n)$

$$= \alpha(w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n)$$
$$= \alpha\langle \mathbf{u}, \mathbf{v} \rangle$$

Hence $\quad \langle \alpha\mathbf{u}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle = w_1u_1^2 + w_2u_2^2 + \cdots + w_nu_n^2 \geq 0$

since $w_1, w_2, \ldots, w_n$ are positive real numbers

If $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, then $w_1u_1^2 + w_2u_2^2 + \cdots + w_nu_n^2 = 0$

i.e. $\quad u_1 = 0, u_2 = 0, \ldots, u_n = 0 \quad$ since $w_i > 0$ for each i, $1 \leq i \leq n$

i.e. $\quad \mathbf{u} = \mathbf{0}$

Hence $\quad \langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Since, $\langle \mathbf{u}, \mathbf{v} \rangle$ has satisfied all the axioms of inner product, R^n is an inner product space with weighted inner product.

EXAMPLE 4.8 In the vector space $C[a, b]$, the function $\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b w(t)f(t)g(t)dt$, is called the weighted inner product, where W is a fixed positive function in $C[a, b]$ and called the *weight function*. Show that the given function is an inner product on $C[a, b]$.

Solution: The axioms of inner product for the given function $\langle \mathbf{f}, \mathbf{g} \rangle$ are as follows:

(i) By using the properties of integration,

$$\langle \mathbf{g}, \mathbf{f} \rangle = \int_a^b w(t)g(t)f(t)dt = \int_a^b w(t)f(t)g(t)dt = \langle \mathbf{f}, \mathbf{g} \rangle$$

Thus, $\quad \langle \mathbf{f}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{f} \rangle.$

(ii) Let $h \in C[a, b]$

$$
\begin{aligned}
\langle \mathbf{f}+\mathbf{g}, \mathbf{h} \rangle &= \int_a^b w(t)\,(f+g)\,(t)h(t)dt \\
&= \int_a^b (w(t)f(t)h(t) + w(t)g(t)h(t))dt \\
&= \int_a^b w(t)f(t)h(t)dt + \int_a^b w(t)g(t)h(t)dt \\
&= \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle
\end{aligned}
$$

Thus, $\langle \mathbf{f}+\mathbf{g}, \mathbf{h} \rangle = \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle$

(iii) For any scalar α, $\langle \alpha\mathbf{f}, \mathbf{g} \rangle = \int_a^b w(t)\,(\alpha f)\,(t)g(t)dt$

$$
\begin{aligned}
&= \alpha \int_a^b w(t)f(t)g(t)dt \\
&= \alpha \langle \mathbf{f}, \mathbf{g} \rangle
\end{aligned}
$$

Thus, $\langle \alpha\mathbf{f}, \mathbf{g} \rangle = \alpha \langle \mathbf{f}, \mathbf{g} \rangle$.

(iv) $\langle \mathbf{f}, \mathbf{f} \rangle = \int_a^b w(t)f^2(t) \geq 0$

since $w(t) > 0$ and $f^2(t) \geq 0$ for all $t \in C[a, b]$, $w(t)f^2(t) \geq 0$ for all $t \in C[a, b]$.

So, $\langle \mathbf{f}, \mathbf{f} \rangle = 0 \Leftrightarrow \int_a^b w(t)f^2(t)dt = 0$

$$
\begin{aligned}
&\Leftrightarrow w(t)f^2(t) = 0 \\
&\Leftrightarrow f(t) = 0 \qquad \text{since } w(t) > 0,\ \forall\ t \in C[a, b],
\end{aligned}
$$

Thus, $\langle \mathbf{f}, \mathbf{f} \rangle = 0 \Leftrightarrow \mathbf{f} = \mathbf{0}$.

Hence the weighted inner product $\langle \mathbf{f}, \mathbf{g} \rangle$ satisfies all the axioms of the inner product of the vector space $C[a, b]$.

EXAMPLE 4.9 In the vector space R^2, show that the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 9u_2v_2$ satisfies all axioms of the inner product.

Solution: Let $\mathbf{u} = [u_1, u_2]$, $\mathbf{v} = [v_1, v_2]$ and $\mathbf{w} = [w_1, w_2] \in R^2$

(i) $\langle \mathbf{v}, \mathbf{u} \rangle = 4v_1u_1 + 9v_2u_2 = 4u_1v_1 + 9u_2v_2 = \langle \mathbf{u}, \mathbf{v} \rangle$

Thus, $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

(ii) $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = 4(u_1+v_1)w_1 + 9(u_2+v_2)w_2$

$$= 4u_1w_1 + 4v_1w_1 + 9u_2w_2 + 9v_2w_2$$

$$= (4u_1w_1 + 9u_2w_2) + (4v_1w_1 + 9v_2w_2)$$

$$= \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$$

Thus, $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$

(iii) For any scalar α,

$$\langle \alpha\mathbf{u}, \mathbf{v}\rangle = 4(\alpha u_1)v_1 + 9(\alpha u_2)v_2$$

$$= \alpha(4u_1v_1 + 9u_2v_2)$$

$$= \alpha\langle \mathbf{u}, \mathbf{v}\rangle$$

Thus, $\langle \alpha\mathbf{u}, \mathbf{v}\rangle = \alpha\langle \mathbf{u}, \mathbf{v}\rangle$.

(iv) $\langle \mathbf{u}, \mathbf{u}\rangle = 4u_1^2 + 9u_2^2 \geq 0$

If $\langle \mathbf{u}, \mathbf{u}\rangle = 0$

then $4u_1^2 + 9u_2^2 = 0$

i.e. $u_1 = 0,\ u_2 = 0$

or $\mathbf{u} = (0, 0) = \mathbf{0}$

Thus, $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Hence the given weighted Euclidean inner product satisfies all the axioms of the inner product.

Inner Product Generated by Matrix

As we know, every vector $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ of R^n can be considered as column matrix of order $n \times 1$,

that is, $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$

Now here we will define inner product in the form of matrix multiplication especially for the weighted Euclidean inner product.

Definition: Inner Product Generated by Matrix

If $\mathbf{A}$ is an invertible matrix of order $n \times n$ and $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{Au} \cdot \mathbf{Av} \qquad \text{where } \mathbf{u} \text{ and } \mathbf{v} \text{ are vectors of } R^n$$

defines an inner product, which is called the *inner product on R^n generated by* $\mathbf{A}$.

Since the dot product in the form of matrix multiplication is given by $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{u}$,

then, $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{Au} \cdot \mathbf{Av}$

i.e. $$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u} = 4u_1v_1 + 9u_2v_2$$

Hence the given inner product is generated by the matrix $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Remark: In general, the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n$$

on R^n is generated by the matrix

$$\mathbf{A} = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{w_3} & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & \sqrt{w_n} \end{bmatrix}$$

EXAMPLE 4.12 Show that the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 13u_1v_1 - 5u_1v_2 - 5u_2v_1 + 17u_2v_2$ is generated by the matrix $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}$.

Solution: By the definition, inner product in the form of matrix multiplication is

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \mathbf{v}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\,\mathbf{A}\mathbf{u} \\ &= [v_1 \;\; v_2]\left\{\begin{bmatrix} 3 & 2 \\ 1 & -4 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}\right\}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= [v_1 \;\; v_2]\begin{bmatrix} 13 & -5 \\ -5 & 17 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= 13u_1v_1 - 5u_1v_2 - 5u_2v_1 + 17u_2v_2 \end{aligned}$$

Hence the given inner product is generated the matrix **A**.

EXAMPLE 4.13 Find the matrix **A** which generates the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + 3u_2v_2$ on R^2.

Solution: Here the given inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + 3u_2v_2$ on R^2 is a weighted inner product on R^2 with weights $w_1 = 1$ and $w_2 = 3$ since the matrix which generates the weighted inner product on R^2 is in the form $\mathbf{A} = \begin{bmatrix} \sqrt{w_1} & 0 \\ 0 & \sqrt{w_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix}$

It is easy to compute $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\,\mathbf{A}\mathbf{u} = u_1v_1 + 3u_2v_2$.

Hence, the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{bmatrix}$ generates the given inner product.

Theorem 4.1 [Properties of Inner Product Space]
Let V be an inner product space and **u**, **v** and **w** be three vectors of V. Let α be a scalar. Then,

(i) $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$

(ii) $\langle \mathbf{u}, \alpha\mathbf{v}\rangle = a\langle \mathbf{u}, \mathbf{v}\rangle$

(iii) $\langle \mathbf{u}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{u}\rangle = 0.$

EXAMPLE 4.14 For the vectors **u**, **v** and **w** in V, if $\langle \mathbf{u}, \mathbf{v}\rangle = 5, \langle \mathbf{v}, \mathbf{w}\rangle = 4, \langle \mathbf{u}, \mathbf{w}\rangle = -2$, evaluate the following expressions:

(i) $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle$ (ii) $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle$ (iii) $\langle 3\mathbf{u}, \mathbf{v}-\mathbf{w}\rangle$

(iv) $\langle 2\mathbf{u}+\mathbf{w}, 3\mathbf{v}\rangle$

Solution: Here we have given $\langle \mathbf{u}, \mathbf{v}\rangle = 5, \langle \mathbf{v}, \mathbf{w}\rangle = 4, \langle \mathbf{u}, \mathbf{w}\rangle = -2$
By using the results of the above theorem, we have

(i) $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle = 5 - 2 = 3$

(ii) $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle = -2 + 4 = 2$

(iii) $\langle 3\mathbf{u}, \mathbf{v}-\mathbf{w}\rangle = \langle 3\mathbf{u}, \mathbf{v}\rangle + \langle 3\mathbf{u}, -\mathbf{w}\rangle = 3\langle \mathbf{u}, \mathbf{v}\rangle - 3\langle \mathbf{u}, \mathbf{w}\rangle = 3(5) - 3(-2) = 21$

(iv) $\langle 2\mathbf{u}+\mathbf{w}, 3\mathbf{v}\rangle = \langle 2\mathbf{u}, 3\mathbf{v}\rangle + \langle \mathbf{w}, 3\mathbf{v}\rangle = 6\langle \mathbf{u}, \mathbf{v}\rangle + 3\langle \mathbf{w}, \mathbf{v}\rangle$

$= 6\langle \mathbf{u}, \mathbf{v}\rangle + 3\langle \mathbf{v}, \mathbf{w}\rangle = 6(5) + 3(4) = 42$

EXAMPLE 4.15 Show that $\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + u_2v_3 + u_3v_2 + u_4v_4$, where $\mathbf{u} = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$, is not an inner product on M_{22} .

Solution: For the given inner product $\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + u_2v_3 + u_3v_2 + u_4v_4$, the first three axioms are easily satisfied. Now, we look at axiom (iv) of inner product. Now,

$$\langle \mathbf{u}, \mathbf{u}\rangle = u_1^2 + 2u_2u_3 + u_4^2$$

We take $\mathbf{u} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$, that is, $u_1 = 1$, $u_2 = 2$, $u_3 = -1$, $u_4 = 1$,

then, $$\langle \mathbf{u}, \mathbf{u}\rangle = 1 - 4 + 1 = -2 < 0$$

Thus the fourth axioms, that is $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ of the inner product is not satisfied by the given function. Hence the given $\langle \mathbf{u}, \mathbf{v}\rangle$ is not an inner product on M_{22}.

EXERCISE SET 1

1. In the vector space R^3, show that the function $\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + u_2v_2 + u_3v_3$ is an inner product on R^3 where $\mathbf{u} = [u_1, u_2, u_3]$ and $\mathbf{v} = [v_1, v_2, v_3]$.

2. Let $\mathbf{u} = [u_1, u_2]$ and $\mathbf{v} = [v_1, v_2]$. Determine which of the following are inner products on R^2. For those which are not, list the axioms that do not hold.

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = 6u_1v_1 + 4u_2v_2$ (ii) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1^2v_1^2 + u_2^2v_2^2$

(iii) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 - u_1v_2 - u_2v_1 + 2u_2v_2$ (iv) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_1v_2 + u_2v_1 + 5u_2v_2$

3. Let $\mathbf{u} = [u_1, u_2, u_3]$ and $\mathbf{v} = [v_1, v_2, v_3]$. Determine which of the following are inner products on R^3. For those which are not, list the axioms that do not hold.

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + 4u_2v_2 + 8u_3v_3$ (ii) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_1v_2$

4. Consider the vector space M_{22} of all 2×2 matrices. If the function $\langle \mathbf{A}, \mathbf{B} \rangle$ is defined by

$$\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) \text{ where } \mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \in M_{22},$$

then show that M_{22} is an inner product space with inner product $\langle \mathbf{A}, \mathbf{B} \rangle$.

5. In the real vector space P_n of all polynomials over R of degree at most n, show that the function

$\langle \mathbf{p}, \mathbf{q} \rangle = \int_0^1 p(t)q(t)\,dt$ is an inner product on P_n where $p, q \in P_n$

6. Show that the Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ on the vector space R^3 is generated by the identity matrix $\mathbf{I}_3$.

7. Show that the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 64u_2v_2 + 5u_3v_3$ on R^3 is generated by

the matrix $\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix}$.

8. Find the matrix $\mathbf{A}$ which generates the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 25u_2v_2 + 8u_3v_3$ on R^3.

9. For the vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in V, if $\langle \mathbf{u}, \mathbf{v} \rangle = -6, \langle \mathbf{v}, \mathbf{w} \rangle = 2, \langle \mathbf{u}, \mathbf{w} \rangle = 9,$ then evaluate the following expressions

(i) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle$ (ii) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle$ (iii) $\langle 2\mathbf{u}, \mathbf{v} + \mathbf{w} \rangle$ (iv) $\langle \mathbf{u} - \mathbf{w}, 4\mathbf{v} \rangle$.

4.2 NORM, DISTANCE AND ANGLE

In Section 4.1, we introduced inner product on the vector space. Now, in this section we will define some geometrical terms like, norm, distance, angle, etc. on an inner product space with respect to an inner product. We will also see some geometrical figures relative to different inner products.

Norm and Distance

In the first course of mathematics – I, Calculus, we defined norm and distance in R^n. Geometrically, the term *norm* of any vector $\mathbf{u}$ is the length of the position vector of $\mathbf{u}$, or in other words, it is the distance of a point $\mathbf{u}$ from the origin.

Distance: $\quad d(\mathbf{u}, \mathbf{v}) = \sqrt{4(u_1 - v_1)^2 + 9(u_2 - v_2)^2} = \sqrt{4(1-0)^2 + 9(0-1)^2}$

$$= \sqrt{13}$$

EXAMPLE 4.18 Find the formulas of norm and distance in M_{22} with respect to an inner product

$$\langle \mathbf{A}, \mathbf{B} \rangle = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 \qquad \text{where } \mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}.$$

Solution: Since $\mathbf{A} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ are two vectors of M_{22}, the formulas of norm and distance with respect to the given inner product in M_{22} are as follows:

Norm: $\quad \|\mathbf{A}\| = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{a_1^2 + a_2^2 + a_3^2 + a_4^2}$

Distance: $\quad d(\mathbf{A}, \mathbf{B}) = \|\mathbf{A} - \mathbf{B}\|$

$$= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 + (a_4 - b_4)^2}$$

since $\quad \mathbf{A} - \mathbf{B} = \begin{bmatrix} a_1 - b_1 & a_2 - b_2 \\ a_3 - b_3 & a_4 - b_4 \end{bmatrix}$

EXAMPLE 4.19 For an inner product space P_2 with an inner product

$$\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2 \qquad \text{where } \mathbf{p} = p_0 + p_1x + p_2x^2,\ \mathbf{q} = q_0 + q_1x + q_2x^2,$$

find the formulas of norm and distance in P_2. Hence find the norms of $1 - x^2$ and $1 - x + x^2$ and the distance between them.

Solution: Since $\mathbf{p} = p_0 + p_1x + p_2x^2$, $\mathbf{q} = q_0 + q_1x + q_2x^2 \in P_2$.

Norm: $\quad \|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{p_0^2 + p_1^2 + p_2^2}$

Distance: $\quad d(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|$

$$= \sqrt{(p_0 - q_0)^2 + (p_1 - q_1)^2 + (p_2 - q_2)^2}$$

since $\quad \mathbf{p} - \mathbf{q} = (p_0 - q_0) + (p_1 - q_1)x + (p_2 - q_2)x^2$

For $\mathbf{p} = 1 - x^2$, $\mathbf{q} = 1 - x + x^2$,

$$\|\mathbf{p}\| = \sqrt{p_0^2 + p_1^2 + p_2^2} = \sqrt{1^2 + 0^2 + (-1)^2} = \sqrt{2}$$

$$\|\mathbf{q}\| = \sqrt{q_0^2 + q_1^2 + q_2^2} = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}$$

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_0 - q_0)^2 + (p_1 - q_1)^2 + (p_2 - q_2)^2} = \sqrt{(1-1)^2 + (0+1)^2 + (-1-1)^2} = \sqrt{5}$$

$$= \int_{-1}^{1} (1 - x + x^2)^2 dx = \int_{-1}^{1} (1 + x^2 + x^4 - 2x - 2x^3 + 2x^2) dx$$

$$= \int_{-1}^{1} (1 + 3x^2 + x^4) dx - \int_{-1}^{1} (2x + 2x^3) dx$$

$$= 2\int_{0}^{1} (1 + 3x^2 + x^4) dx = 2\left(x + x^3 + \frac{x^5}{5} \right)\Bigg|_0^1$$

$$= 2\left(1 + 1 + \frac{1}{5}\right) = \frac{22}{5}$$

Hence $\qquad \|\mathbf{q}\| = \sqrt{\dfrac{22}{5}}$

Distance: Here $p(x) - q(x) = x - 2x^2$

$$d(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\| = \sqrt{\int_{-1}^{1} (x - 2x^2)^2 dx}$$

$$= \sqrt{\int_{-1}^{1} (x^2 - 4x^3 + 4x^4) dx}$$

$$= \sqrt{2\int_{0}^{1} (x^2 + 4x^4) dx}$$

$$= \sqrt{2\left(\frac{x^3}{3} + 4\,\frac{x^5}{5}\right)_0^1}$$

$$= \sqrt{2\left(\frac{1}{3} + \frac{4}{5}\right)}$$

$$= \sqrt{\frac{34}{15}}$$

Theorem 4.2 [Properties of Norm and Distance]

Let V be an inner product space. Then for arbitrary vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{z}$ in V and scalar α

(i) $\|\alpha \mathbf{u}\| = |\alpha|\,\|\mathbf{u}\|$

(ii) $\|\mathbf{u}\| \geq 0$ and $\|\mathbf{u}\| = 0$ if $\mathbf{u} = \mathbf{0}$

(iii) Cauchy–Schwarz inequality $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$

(iv) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

(v) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

(vi) $d(\mathbf{u}, \mathbf{v}) \geq 0$ and $d(\mathbf{u}, \mathbf{v}) = 0$ if $\mathbf{u} = \mathbf{v}$

(vii) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{z}) + d(\mathbf{z}, \mathbf{v})$

EXAMPLE 4.21 Verify the Cauchy–Schwarz inequality for the vectors $1 - x^2$ and $1 - x + x^2$ in P_2 with an inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} p(x)q(x)dx$.

Solution: The Cauchy–Schwarz inequality for **p** and **q** is $|\langle \mathbf{p}, \mathbf{q} \rangle| \le \|\mathbf{p}\|\,\|\mathbf{q}\|$. To verify this inequality we need the values of $\langle \mathbf{p}, \mathbf{q} \rangle$, $\|\mathbf{p}\|$ and $\|\mathbf{q}\|$ which are calculated in Example 4.20.

$$\|\mathbf{p}\| = \frac{4}{\sqrt{15}}, \quad \|\mathbf{q}\| = \sqrt{\frac{22}{5}}$$

$$\begin{aligned}
\langle \mathbf{p}, \mathbf{q} \rangle &= \int_{-1}^{1} p(x)q(x)dx \\
&= \int_{-1}^{1} (1-x^2)(1-x+x^2)dx \\
&= \int_{-1}^{1} (1-x+x^3-x^4)dx \\
&= 2\int_{0}^{1} (1-x^4)dx \\
&= 2\left(x-\frac{x^5}{5}\right)\Bigg|_0^1 \\
&= 2\left(1-\frac{1}{5}\right) \\
&= \frac{8}{5}
\end{aligned}$$

$$\|\mathbf{p}\|\,\|\mathbf{q}\| = \frac{4}{\sqrt{15}} \cdot \sqrt{\frac{22}{5}} = \frac{4}{5} \cdot \sqrt{\frac{22}{3}}$$

since

$$2 \le \sqrt{\frac{22}{5}}$$

$$\frac{8}{5} \le \frac{4}{5}\sqrt{\frac{22}{5}}$$

Hence

$$|\langle \mathbf{p}, \mathbf{q} \rangle| \le \|\mathbf{p}\|\,\|\mathbf{q}\|$$

Definition: Angle between Two Vectors

If θ is the angle between the vectors **u** and **v** in inner product space V, then $\theta = \cos^{-1}\left(\dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$.

EXAMPLE 4.22 Find the angle between the vectors $\mathbf{u} = [1, 2]$ and $\mathbf{v} = [-2, 1]$ in R^2 with respect to an inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 9u_2v_2$.

Solution: For the given $\mathbf{u} = [1, 2]$ and $\mathbf{v} = [-2, 1]$

$$\langle \mathbf{u}, \mathbf{v} \rangle = 4(1)(-2) + 9(2)(1) = 10$$

$$\|\mathbf{u}\| = \sqrt{4(1)^2 + 9(2)^2} = \sqrt{40} = 2\sqrt{10}$$

$$\|\mathbf{v}\| = \sqrt{4(-2)^2 + 9(1)^2} = \sqrt{25} = 5$$

If θ is the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$, then

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$$

$$= \cos^{-1}\left(\frac{10}{(2\sqrt{10})(5)}\right)$$

$$= \cos^{-1}\left(\frac{1}{\sqrt{10}}\right)$$

EXAMPLE 4.23 If $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, find the angle between $\mathbf{u}$ and $\mathbf{v}$ in M_{22} with respect to inner product $\langle \mathbf{A}, \mathbf{B} \rangle = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$.

Solution: For $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$$\langle \mathbf{A}, \mathbf{B} \rangle = (1)(0) + (0)(1) + (0)(1) + (1)(0) = 0$$

If θ is the angle between the vectors $\mathbf{u}$ and $\mathbf{v}$, then

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{A}, \mathbf{B} \rangle}{\|\mathbf{A}\|\|\mathbf{B}\|}\right)$$

$$= \cos^{-1}(0) \qquad \text{since } \|\mathbf{A}\| = 1 \text{ and } \|\mathbf{B}\| = 1$$

$$\therefore \qquad \theta = \frac{\pi}{2}$$

Definition: Unit Sphere

The set of vectors $\mathbf{u}$ in an inner product space V such that $\|\mathbf{u}\| = 1$ is called a *unit sphere* in V.

EXAMPLE 4.24 Find the equation of unit sphere

(i) in R^2 with an inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_2v_2$

(ii) in R^2 with an inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 9u_2v_2$

EXERCISE SET 2

1. For each of the following compute $\langle \mathbf{u}, \mathbf{v} \rangle, \|\mathbf{u}\|$ and $d(\mathbf{u}, \mathbf{v})$ for the given pair of vectors and inner product.
   (i) $\mathbf{u} = [2, -1, 4]$ and $\mathbf{v} = [3, 2, 0]$ in R^3 with standard Euclidean inner product.
   (ii) $\mathbf{u} = [2, -1, 4]$ and $\mathbf{v} = [3, 2, 0]$ in R^3 with inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + 6u_2v_2 + \frac{1}{5}u_3v_3$
   (iii) $\mathbf{u} = x$ and $\mathbf{v} = x^2$ in $C[0, 1]$ using the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \int_0^1 uv\, dx$.
2. Verify the Cauchy–Schwarz inequality for the vectors $\mathbf{A} = \begin{bmatrix} 0 & -2 \\ 2 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$ in R^2 with respect to inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B})$.
3. Verify the Cauchy–Schwarz inequality for the vectors $-4 + 2x + x^2$ and $8 - 4x - 2x^2$ in P_2 with respect to inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \int_0^1 p(x)q(x)dx$.
4. Find the angle between the vectors $\mathbf{u} = [1, 2, 3]$ and $\mathbf{v} = [1, 1, 1]$ in R^3 with respect to inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_2v_2 + u_3v_3$.
5. Find the angle between the vectors $\mathbf{A} = \begin{bmatrix} 2 & 4 \\ -1 & 3 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -3 & 1 \\ 4 & 2 \end{bmatrix}$ in R^2 with respect to inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B})$.
6. Find the equations of unit sphere R^3 with inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 9u_2v_2 + 16u_3v_3$.

4.3 ORTHOGONAL VECTORS, ORTHOGONAL COMPLEMENTS

Orthogonal Vectors

In this section, we will define orthogonal vectors in the inner product spaces. But, we know that two vectors are orthogonal (or perpendicular) if the angle between them is $\pi/2$. In the previous section, we defined the angle between two vectors in the inner product space V by the formula $\cos\theta = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$ where θ is an angle between the vectors $\mathbf{u}$ and $\mathbf{v}$. So, $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (that is, $\theta = \pi/2$). If $\cos\theta = 0$, then $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. The precise definition is given below.

Definition:* *Orthogonal Vectors

If $\mathbf{u}$ and $\mathbf{v}$ are two vectors in an inner product space V, then they are said to be orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Solution: For the given vectors, $\mathbf{u} = [1, 2]$ and $\mathbf{v} = [-8, 9]$

Euclidean inner product:
$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle &= u_1v_1 + u_2v_2\\ &= (1)(-8) + (2)(9)\\ &= 10\\ &\neq 0 \qquad \text{(i)}\end{aligned}$$

Given inner product
$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle &= 9u_1v_1 + 4u_2v_2\\ &= 9(1)(-8) + 4(2)(9)\\ &= -72 + 72 = 0 \qquad \text{(ii)}\end{aligned}$$

Hence, from the result (i) and (ii), the given pair of vectors is orthogonal with respect to the given weighted inner product but not orthogonal with respect to Euclidean inner product in R^2.

EXAMPLE 4.27 Find the value of k, for which the following pairs of vectors are orthogonal with respect to Euclidean inner product.

(i) $\mathbf{u} = [-1, 2, 3]$ and $\mathbf{v} = [k, -4, k]$
(ii) $\mathbf{u} = [k, 3, k]$ and $\mathbf{v} = [-2, 5, k]$

Solution: To find the value of k for the orthogonal vectors, we need to solve the equation $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

(i) Here $\mathbf{u} = [-1, 2, 3]$ and $\mathbf{v} = [k, -4, k]$
Suppose $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors with respect to Euclidean inner product, then

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

or $$u_1v_1 + u_2v_2 + u_3v_3 = 0$$

or $$(-1)(k) + (2)(-4) + (3)(k) = 0$$

or $$-k - 8 + 3k = 2k - 8 = 0$$

$\therefore$ $$k = 4$$

Therefore, for $k = 4$, the given vectors are orthogonal vectors.

(ii) If the given vectors $\mathbf{u} = [k, 3, k]$ and $\mathbf{v} = [-2, 5, k]$ are orthogonal with respect to Euclidean inner product, then

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

or $$u_1v_1 + u_2v_2 + u_3v_3 = 0$$

or $$(k)(-2) + (3)(5) + (k)(k) = 0$$

or $$(-2k) + 15 + k^2 = 0$$

or $$k^2 - 2k + 15 = 0$$

or $$(k - 5)(k + 3) = 0$$

$\therefore$ $$k = 5 \text{ or } k = -3$$

Therefore, the given vectors are orthogonal for $k = 5$ or -3.

EXAMPLE 4.28 Find all vectors $\mathbf{u} = [u_1, u_2]$, orthogonal to $[2, -1]$ with respect to Euclidean inner product.

Solution: We have given $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 3 & 2 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 6 \\ 4 & -3 \end{bmatrix}$

$$\begin{aligned} \langle \mathbf{A}, \mathbf{B} \rangle &= a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 \\ &= (0)(1) + (-1)(6) + (3)(4) + (2)(-3) \\ &= 0 - 6 + 12 - 6 \\ &= 0 \end{aligned}$$

That is $\langle \mathbf{A}, \mathbf{B} \rangle = 0$

Therefore **A** and **B** are orthogonal matrices.

EXAMPLE 4.31 Check the orthogonality of the vectors $\mathbf{p}(x) = 1 - 3x - x^2$, $\mathbf{q}(x) = 5 + 3x - 4x^2$ in the inner product space P_2 with inner product $\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2$ where $\mathbf{p}(x) = p_0 + p_1x + p_2x^2$, $\mathbf{q}(x) = q_0 + q_1x + q_2x^2$.

Solution: For the given vectors $\mathbf{p}(x) = 1 - 3x - x^2$, $\mathbf{q}(x) = 5 + 3x - 4x^2$, we have

$$\begin{aligned} \langle \mathbf{p}, \mathbf{q} \rangle &= (1)(5) + (-3)(3) + (-1)(-4) \\ &= 5 - 9 + 4 \\ &= 0 \end{aligned}$$

That is $\langle \mathbf{p}, \mathbf{q} \rangle = 0$. Therefore, $\mathbf{p}(x)$ and $\mathbf{q}(x)$ are orthogonal vectors in P_2.

EXAMPLE 4.32 Check the orthogonality of the vectors $\mathbf{p}(x) = x^2 - 1$, $\mathbf{q}(x) = x$ in the inner product space P_{2s} with inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} \mathbf{p}(x)\, \mathbf{q}(x) dx$

Solution: Here $\mathbf{p}(x) = x^2 - 1$, $\mathbf{q}(x) = x$

$$\begin{aligned} \langle \mathbf{p}, \mathbf{q} \rangle &= \int_{-1}^{1} (x^2 - 1) x dx \\ &= \int_{-1}^{1} (x^3 - x) dx \\ &= 0 \end{aligned}$$

Hence $\langle \mathbf{p}, \mathbf{q} \rangle = 0$

Therefore the given vectors are orthogonal in P_2 .

EXAMPLE 4.33 Show that $\mathbf{f}(x) = \cos 2\pi x - \sin 2\pi x$ and $g(x) = \cos 2\pi x + \sin 2\pi x$ are orthogonal vector in $C[0, 1]$ with inner product $\langle \mathbf{f}, \mathbf{g} \rangle = \int_{0}^{1} \mathbf{f}(x)\, \mathbf{g}(x) dx$.

Solution: For the given functions, $\mathbf{f}(x) = \cos 2\pi x - \sin 2\pi x$, $\mathbf{g}(x) = \cos 2\pi x + \sin 2\pi x$

$$
\begin{aligned}
\langle \mathbf{f}, \mathbf{g} \rangle &= \int_0^1 \mathbf{f}(x)\,\mathbf{g}(x)\,dx \\
&= \int_0^1 (\cos^2 2\pi x - \sin^2 2\pi x)\,dx \\
&= \int_0^1 (\cos 4\pi x)\,dx \\
&= \left(\frac{\sin 4\pi x}{4\pi}\right)_0^1 \\
&= \frac{1}{4\pi}(\sin 4\pi - \sin 0) = 0
\end{aligned}
$$

That is $\langle \mathbf{f}, \mathbf{g} \rangle = 0.$ Therefore, **f** and **g** are orthogonal vectors in $C[0, 1]$.

Theorem 4.3 [Pythagoras Theorem]

If two vectors **u** and **v** are orthogonal in an inner product space, then $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

EXAMPLE 4.34 Verify Pythagoras theorem for vectors $\mathbf{u} = [1, 1, 0]$ and $\mathbf{v} = [1, -1, 1]$ in the Euclidean inner product space R^3.

Solution: Here $\mathbf{u} = [1, 1, 0]$, $\mathbf{v} = [1, -1, 1]$

Euclidean inner product
$$
\begin{aligned}
\langle \mathbf{u}, \mathbf{v} \rangle &= u_1v_1 + u_2v_2 + u_3v_3 \\
&= (1)(1) + (1)(-1) + (0)(1) \\
&= 0
\end{aligned}
$$

Therefore the given vectors **u** and **v** are orthogonal vectors.

$$
\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 + u_2^2 + u_3^2 = 1 + 1 = 0 = 2
$$

$$
\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle = v_1^2 + v_2^2 + v_3^2 = 1 + 1 + 1 = 3
$$

$$
\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = 2 + 3 = 5 \tag{i}
$$

$$
\mathbf{u} + \mathbf{v} = [1, 1, 0] + [1, -1, 1] = [2, 0, 1]
$$

$$
\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = 4 + 0 + 1 = 5 \tag{ii}
$$

From the results (i) and (ii),

$$
\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2
$$

Hence the Pythogoras theorem is verified.

$$
\begin{aligned}
\|\mathbf{p}\|^2 &= \langle \mathbf{p}, \mathbf{p} \rangle \\
&= \int_{-1}^{1} \mathbf{p}(x)\mathbf{p}(x)dx \\
&= \int_{-1}^{1} (x^2-1)^2 dx \\
&= \int_{-1}^{1} (x^4 - 2x^2 + 1)dx \\
&= \int_{0}^{1} (x^4 - 2x^2 + 1)dx \\
&= 2\left(\frac{x^5}{5} - \frac{2}{3}x^3 + x\right)_0^1 \\
&= 2\left(\frac{1}{5} - \frac{2}{3} + 1\right) \\
&= 2\left(\frac{3-10+15}{15}\right) \\
&= \frac{16}{15} \\
\|\mathbf{q}\|^2 &= \langle \mathbf{q}, \mathbf{q} \rangle = \int_{-1}^{1} \mathbf{q}(x)\mathbf{p}(x)dx \\
&= \int_{-1}^{1} x^2 dx \\
&= 2\int_{0}^{1} x^2 dx \\
&= 2\left(\frac{x^3}{3}\right)_0^1 \\
&= \frac{2}{3} \\
\|\mathbf{p}\|^2 + \|\mathbf{q}\|^2 &= \frac{16}{15} + \frac{2}{3} = \frac{26}{15} \qquad \text{(i)} \\
(\mathbf{p} + \mathbf{q})x &= \mathbf{p}(x) + \mathbf{q}(x) = x^2 + x - 1
\end{aligned}
$$

$$\begin{aligned}\|\mathbf{p}+\mathbf{q}\|^2 &= \langle \mathbf{p}+\mathbf{q}, \mathbf{p}+\mathbf{q}\rangle = \int_{-1}^{1}(x^2+x-1)^2\,dx\\ &= \int_{-1}^{1}(x^4+x^2+1+2x^3-2x^2-2x)dx\\ &= 2\int_0^1(x^4-x^2+1)dx\\ &= 2\left(\frac{x^5}{5}-\frac{x^3}{3}+x\right)_0^1\\ &= 2\left(\frac{1}{5}-\frac{1}{3}+1\right)\\ &= \frac{26}{15}\end{aligned}$$

From the result (i) and (ii),

$$\|\mathbf{p}\|^2+\|\mathbf{q}\|^2 = \|\mathbf{p}+\mathbf{q}\|^2$$

Hence the Pythagoras theorem is verified.

EXAMPLE 4.38 Verify the Pythagoras theorem for the function $\mathbf{f}(x) = \cos 2\pi x - \sin 2\pi x$ and $\mathbf{g}(x) = \cos 2\pi x + \sin 2\pi x$ in $C[0, 1]$ with inner product $\langle \mathbf{f}, \mathbf{g}\rangle = \int_0^1 \mathbf{f}(x)\mathbf{g}(x)dx$.

Solution: We showed in Example 4.33 that **f** and **g** are orthogonal vectors in $C[0, 1]$.

$$\begin{aligned}\|\mathbf{f}\|^2 &= \langle \mathbf{f}, \mathbf{f}\rangle\\ &= \int_0^1(\cos 2\pi x-\sin 2\pi x)^2\,dx\\ &= \int_0^1(1-2\sin 2\pi x\cos 2\pi x)dx\\ &= \int_0^1(1-\sin 4\pi x)dx\\ &= \left(x+\frac{\cos 4\pi x}{4\pi}\right)_0^1\\ &= \left(1+\frac{1}{4\pi}-0-\frac{1}{4\pi}\right) = 1\end{aligned}$$

EXAMPLE 4.39 Show that the vector $\mathbf{u} = [1, 1, 0]$ is orthogonal to set $W = \{[0, 0, 1], [1, -1, 1]\}$ in R^3.

Solution: Let $\mathbf{w}_1 = [0, 0, 1]$, $\mathbf{w}_2 = [1, -1, 1]$ in W. We want to prove that $\mathbf{u} = [1, 1, 0]$ is orthogonal to $W = \{\mathbf{w}_1, \mathbf{w}_2\}$. It is sufficient to show that $\langle \mathbf{u}, \mathbf{w}_1 \rangle = 0, \langle \mathbf{u}, \mathbf{w}_2 \rangle = 0$.

$$\langle \mathbf{u}, \mathbf{w}_1 \rangle = (1)(0) + (1)(0) + (0)(1) = 0$$

$$\langle \mathbf{u}, \mathbf{w}_2 \rangle = (1)(1) + (1)(-1) + (0)(1) = 0$$

Therefore $\mathbf{u}$ is orthogonal to every vector in W. Hence $\mathbf{u}$ is orthogonal to the set W.

Definition: Orthogonal Complement

Let W be a subspace of an inner product space V. The *orthogonal complement* of W is the set of all vectors which are orthogonal to the subspace W. It is denoted by $W^{\perp}$ [read as "W perp"].

In the above definition, W is a subspace of V, therefore $\mathbf{0} \in W$. Also the zero vector $\mathbf{0}$ is orthogonal to every vector in W, therefore it is also an element of $W^{\perp}$. So, $\mathbf{0}$ is the only common vector in W and $W^{\perp}$. Some more properties of $W^{\perp}$ are given in the following theorem.

Theorem 4.4 [Properties of $W^{\perp}$]

Let W be a subspace of an inner product space V. If $W^{\perp}$ is an orthogonal complement of W, then

(i) a vector $\mathbf{u}$ in $W^{\perp}$ if and only if $\mathbf{u}$ is orthogonal to every vector in the subspace W.

(ii) $\mathbf{0}$ is the only common vector in W and $W^{\perp}$, that is, $W \cap W^{\perp} = \{0\}$

(iii) $W^{\perp}$ is a subspace of V.

(iv) The orthogonal complement of orthogonal complement $[W^{\perp}]$ of W is W itself, that is $[W^{\perp}]^{\perp} = W$.

(v) The orthogonal complement of vector space V is $\{0\}$, that is $V^{\perp} = \{0\}$ and also $\{0\}^{\perp} = V$.

Sometimes the following theorem is called the "Fundamental Theorem of Linear Algebra".

Theorem 4.5

If $\mathbf{A}$ is a matrix of order $m \times n$, then $\mathbf{A}$ has orthogonal complements in R^n.

(i) The row space of $\mathbf{A}$ and the null space of $\mathbf{A}$ are orthogonal complements in R^n, that is, $N_A = R_A^{\perp}$

(ii) The column space of $\mathbf{A}$ and the null space of $\mathbf{A}^{\mathrm{T}}$ are orthogonal complements in R^m, that is, $N_{A^T} = C_A^{\perp}$

Remark: Recall the following facts. If $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, then

$$R_A = \{\alpha_1[a_{11}, a_{12}, \ldots, a_{1n}] + \alpha_2[a_{21}, a_{22}, \ldots, a_{2n}] + \cdots + \alpha_m[a_{m1}, a_{m2}, \ldots, a_{mn}] \mid \alpha_1, \alpha_2, \ldots, \alpha_m \in R\} \in R^n$$

$$C_A = \{\beta_1[a_{11}, a_{21}, \ldots, a_{m1}] + \beta_2[a_{12}, a_{22}, \ldots, a_{m2}] + \cdots + \beta_n[a_{1n}, a_{2n}, \ldots, a_{mn}] \mid \beta_1, \beta_2, \ldots, \beta_n \in R\} \in R^m$$

$$N_A = \{\mathbf{x} = [x_1, x_2, \ldots, x_n] \mid \mathbf{Ax} = \mathbf{0}\} \in R^n$$

$$N_{A^T} = \{\mathbf{y} = [y_1, y_2, \ldots, y_m] \mid \mathbf{A}^{\mathrm{T}}\mathbf{y} = \mathbf{0}\} \in R^m$$

EXAMPLE 4.40 Find a basis for the orthogonal complement of the subspace of R^3 spanned by the vectors $\mathbf{v}_1 = [1, -1, 3]$, $\mathbf{v}_2 = [5, -4, -4]$, $\mathbf{v}_3 = [7, -6, 2]$.

Solution: Let W be a subspace of R^3 spanned by the given vector $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

The space W is the same as the row space of the matrix,

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$$

From Theorem 4.5, null space of matrix **A** is orthogonal complement of row space of **A**, that is, the null space of matrix **A** is orthogonal complement of the space W. Therefore, now we will try to find a basis of the null space of matrix **A**. Since the null space **A** is a solution space of the equation $\mathbf{Ax} = \mathbf{0}$, we have

$$\begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} x_1 - x_2 + 3x_3 &= 0 \\ 5x_1 - 4x_2 - 4x_3 &= 0 \\ 7x_1 - 6x_2 + 2x_3 &= 0 \end{aligned}$$

To find the nonzero solution of the above system of linear equations, we use the Gauss-elimination method. The augmented matrix for the above system is

$$\begin{bmatrix} 1 & -1 & 3 & : & 0 \\ 5 & -4 & -4 & : & 0 \\ 7 & -6 & 2 & : & 0 \end{bmatrix}$$ by applying the row operations $R_2 - 5R_1$, $R_3 - 7R_1$

$$\begin{bmatrix} 1 & -1 & 3 & : & 0 \\ 0 & 1 & -19 & : & 0 \\ 0 & 1 & -19 & : & 0 \end{bmatrix}$$ by applying the row operations $R_3 - R_2$, $R_1 + R_2$

$$\begin{bmatrix} 1 & 0 & -16 & : & 0 \\ 0 & 1 & -19 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Therefore we get $x_1 - 16x_3 = 0$ and $x_2 - 19x_3 = 0$

Hence $\qquad x_1 = 16x_3$ and $x_2 = 19x_3$

$$[x_1, x_2, x_3] = [16x_3, 19x_3, x_3]$$

$\therefore \qquad [x_1, x_2, x_3] = x_3[16, 19, 1]$

Therefore, the null space $N_A = \{k\,[16, 19, 1] \mid k \in R\}$

Hence $\{[16, 19, 1]\}$ is a basis for the orthogonal complement of the vector space spanned by the given vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

In other words, W is a null space of matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \end{bmatrix}$

From Theorem 4.5, the row space R^n of matrix **A** is an orthogonal complement of null space of matrix **A**. Therefore, the row space R_A of matrix **A** is an orthogonal complement of the given space W.

$$\therefore \qquad W^{\perp} = R_A = \{\alpha[1, 2, 1] + \beta[1, -2, 1] \mid \alpha, \beta \in R\}$$
$$= \{(\alpha + \beta, 2\alpha - 2\beta, \alpha + \beta \mid \alpha, \beta \in R\}$$

Therefore, the parametric equations of $W^{\perp}$ are $x = \alpha + \beta$, $y = 2\alpha - 2\beta$, $z = \alpha + \beta$

$$2x + y = 2\alpha, \; 2x - y = 4\beta$$
$$2z + y = 2\alpha, \; 2z - y = 4\beta$$
$$2x + y = 2z + y, \; 2x - y = 2z - y$$
$$\therefore \qquad x = z$$
$$\text{or} \qquad x - z = 0$$

Hence $x - z = 0$ is the required equation of $W^{\perp}$.

EXERCISE SET 3

1. Determine which of the following pairs of vectors are orthogonal with respect to Euclidean inner product.

   (i) $\mathbf{u} = [0, 1]$, $\mathbf{v} = [1, -1]$ (ii) $\mathbf{u} = [3, -1, 2]$, $\mathbf{v} = [-1, 5, 4]$

   (iii) $\mathbf{u} = [3, 1, 1, -4]$, $\mathbf{v} = [2, -3, 5, 2]$ (iv) $\mathbf{u} = [1, 0, 0, 0, 2]$, $\mathbf{v} = [0, 3, -7, 9, 1]$

2. Show that the vectors $\mathbf{u} = [1, 1]$ and $\mathbf{v} = [1, -1]$ are orthogonal vectors in R^2 with respect to Euclidean inner product but they are not orthogonal in R^2 with respect to weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + u_2v_2$.

3. Find the values of k, for which the following pairs of vectors are orthogonal with respect to Euclidean inner product.

   (i) $\mathbf{u} = [2, -4]$ and $\mathbf{v} = [k, k - 4]$ (ii) $\mathbf{u} = [k, k, 3]$ and $\mathbf{v} = [1, k, -4]$

4. Find all vectors $\mathbf{u} = [u_1, u_2]$, orthogonal to $[1, -1]$ with respect to weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + 2u_2v_2$.

5. Find all the vectors in R^3 that are orthogonal to $[1, 3, -1]$ with respect to Euclidean inner product. Do they form a vector subspace of R^3?

6. Check the orthogonality between matrices $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix}$ in the inner product space M_{22} with inner product $\langle \mathbf{A}, \mathbf{B} \rangle = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4$.

7. Check the orthogonality of the vectors $\mathbf{p}(x) = 1 + 2x + 3x^2$, $\mathbf{q}(x) = 4 + x - 2x^2$ in the inner product space p_2 with inner product $\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2$ where $\mathbf{p}(x) = p_0 + p_1x + p_2x^2$, $\mathbf{q}(x) = q_0 + q_1x + q_2x^2$.

8. Show that $\mathbf{f}(x) = \sin x - \cos x$ and $\mathbf{g}(x) = \sin x + \cos x$ in $C[0, \pi/2]$ are orthogonal with inner product $\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^{\pi/2} \mathbf{f}(x)\mathbf{g}(x)dx$.

9. Verify the Pythagoras theorem for the function $\mathbf{f}(t) = 1$, $\mathbf{g}(t) = t$ and $\mathbf{h}(t) = 1 + t$ in $C[-1, 1]$, with respect to inner product $\langle \mathbf{f}, \mathbf{g} \rangle = \int_{-1}^{1} \mathbf{f}(t)\mathbf{g}(t)dt$.

10. Show that the vector $\mathbf{u} = [-1, -1, 1, 0]$ is orthogonal to the set $W = \{[1, 4, 5, 2], [2, 1, 3, 0], [-1, 3, 2, 2]\}$ in R^4.

11. Find a basis for the orthogonal complement of the subspace of R^3 spanned by the vectors $\mathbf{v}_1 = [2, 0, -1]$, $\mathbf{v}_2 = [4, 0, -2]$.

12. If the line $y = 3x$, $x \in R$ is a subspace W in R^2, then find an equation for $W^{\perp}$.

13. If W is the intersection of the two planes $x + y + z = 0$ and $x - y + z = 0$, then find an equation for $W^{\perp}$.

4.4 ORTHOGONAL BASIS

In Section 4.3, we defined orthogonal vectors in an inner product space. Now, we will try to extend the concept of orthogonality to the set of vectors in inner product space. Also, we will discuss the orthonormal basis for an inner product space which has some nice properties and is also useful to simplify the problems.

Definition: Orthogonal Set

A set W of vectors $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_p\}$ in an inner product space is said to be an orthogonal set if each pair of distinct vectors from W is orthogonal, that is,

$$\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0 \qquad \text{wherever } i \neq j$$

EXAMPLE 4.43 Show that $\left\{[3, 1, 1], [-1, 2, 1], \left[-\frac{1}{2}, -2, \frac{7}{2}\right]\right\}$ is an orthogonal set in R^3 with Euclidean inner product.

Solution: Let $\mathbf{w}_1 = [3, 1, 1]$, $\mathbf{w}_2 = [-1, 2, 1]$, $\mathbf{w}_3 = \left[-\frac{1}{2}, -2, \frac{7}{2}\right]$ and $W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. To check the orthogonality, we have to show that each pair of distinct vectors is orthogonal. In the Euclidean inner product space R^3,

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 3(-1) + 1(2) + 1(1) = 0$$

$$\langle \mathbf{w}_2, \mathbf{w}_3 \rangle = (-1)\left(-\frac{1}{2}\right) + 2(-2) + (1)\left(\frac{7}{2}\right) = 0$$

$$\langle \mathbf{w}_3, \mathbf{w}_1 \rangle = \left(-\frac{1}{2}\right)(3) + (-2)(1) + \left(\frac{7}{2}\right)(1) = 0$$

Since each pair of distinct vectors of W is orthogonal, W is an orthogonal set in R^3 with Euclidean product.

EXAMPLE 4.44 Show that {[1, 0, –1], [1, 0, 9], [0, 1, 0]} is an orthogonal set in R^3 with weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$.

Solution: Let $\mathbf{w}_1 = [1, 0, -1]$, $\mathbf{w}_2 = [1, 0, 9]$, $\mathbf{w}_3 = [0, 1, 0]$ and $W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$

To check the orthogonality, we have to show that each pair of distinct vectors is orthogonal. For the given weighted inner product,

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 9(1)(1) + 4(0)(0) + (-1)(9) = 9 - 9 = 0$$

$$\langle \mathbf{w}_2, \mathbf{w}_3 \rangle = 9(1)(0) + 4(0)(1) + (9)(0) = 0$$

$$\langle \mathbf{w}_3, \mathbf{w}_1 \rangle = 9(0)(1) + 4(1)(0) + (-1)(0) = 0$$

Since each pair of distinct vectors of W is orthogonal, W is an orthogonal set in R^3 with the given weighted Euclidean inner product.

EXAMPLE 4.45 Check the orthogonality of the set $\left\{\frac{4}{5}x + \frac{3}{5}x^2, -\frac{3}{5}x + \frac{4}{5}x^2, 1\right\}$ in P_2 with an inner product $\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2$ where $\mathbf{p} = p_0 + p_1x + p_2x^2$, $\mathbf{q} = q_0 + q_1x + q_2x^2$.

Solution: Let $\mathbf{p}(x) = \frac{4}{5}x + \frac{3}{5}x^2, \quad \mathbf{q}(x) = -\frac{3}{5}x + \frac{4}{5}x^2, \mathbf{r}(x) = 1$

And $P_2 = \{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$. For the given inner product in P_2,

$$\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2 = (0)(0) + \left(\frac{4}{5}\right)\left(-\frac{3}{5}\right) + \left(\frac{3}{5}\right)\left(\frac{4}{5}\right) = 0$$

$$\langle \mathbf{q}, \mathbf{r} \rangle = (0)(1) + \left(-\frac{3}{5}\right)(0) + \left(\frac{4}{5}\right)(0) = 0$$

$$\langle \mathbf{r}, \mathbf{p} \rangle = 1(0) + (0)\left(\frac{4}{5}\right) + (0)\left(\frac{3}{5}\right) = 0$$

Since each pair in set P_2 is an orthogonal pair, the given set is an orthogonal set in P_2 with the given inner product.

EXAMPLE 4.46 Is the set $W = \{1, x, x^2\}$ an orthogonal set in P_2 with an inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} \mathbf{p}(x)\mathbf{q}(x)dx$?

Solution: Let $\mathbf{p}(x) = 1$, $\mathbf{q}(x) = x$, $\mathbf{r}(x) = x^2$. For the given inner product,

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} (1)(x)dx = \int_{-1}^{1} (x)dx = 0$$

$$\langle \mathbf{q}, \mathbf{r} \rangle = \int_{-1}^{1} (x)(x^2)dx = \int_{-1}^{1} x^3dx = 0$$

$$\langle \mathbf{r}, \mathbf{p} \rangle = \int_{-1}^{1} (x^2)(1)dx = 2\int_{0}^{1} x^2 dx = \frac{2}{3} \neq 0$$

Since $\langle \mathbf{r}, \mathbf{p} \rangle \neq 0$, the set $W = \{1, x, x^2\}$ is not an orthogonal set in P_2 with the above inner product.

EXAMPLE 4.47 Check the orthogonality of the set $\left\{\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} -2 & 2 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ in M_{22}

With inner product $\langle \mathbf{U}, \mathbf{V} \rangle = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$, where $\mathbf{U} = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}, \mathbf{V} = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$.

Solution: Let $\mathbf{V}_1 = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}, \mathbf{V}_2 = \begin{bmatrix} -2 & 2 \\ 3 & 2 \end{bmatrix}, \mathbf{V}_3 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \mathbf{V}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

We have to show that set $S = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$ is an orthogonal set in M_{22} with respect to inner product $\langle \mathbf{U}, \mathbf{V} \rangle = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$

$$\langle \mathbf{V}_1, \mathbf{V}_2 \rangle = (1)(-2) + (-1)(2) + (2)(3) + (-1)(2) = -2 - 2 + 6 - 2 = 0$$

$$\langle \mathbf{V}_1, \mathbf{V}_3 \rangle = (1)(1) + (-1)(2) + (2)(0) + (-1)(-1) = 1 - 2 + 0 + 1 = 0$$

$$\langle \mathbf{V}_1, \mathbf{V}_4 \rangle = (1)(1) + (-1)(0) + (2)(0) + (-1)(1) = 1 + 0 + 0 - 1 = 0$$

$$\langle \mathbf{V}_2, \mathbf{V}_3 \rangle = (-2)(1) + (2)(2) + (3)(0) + (2)(-1) = -2 + 4 + 0 - 2 = 0$$

$$\langle \mathbf{V}_2, \mathbf{V}_4 \rangle = (-2)(1) + (2)(0) + (3)(0) + (2)(1) = -2 + 0 + 0 + 2 = 0$$

$$\langle \mathbf{V}_3, \mathbf{V}_4 \rangle = (1)(1) + (2)(0) + (0)(0) + (-1)(1) = 1 + 0 + 0 - 1 = 0$$

Therefore, $\langle \mathbf{V}_i, \mathbf{V}_j \rangle = 0$ for each $i \neq j, \quad i, j = 1, 2, 3, 4$

Hence the set $S = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$ is an orthogonal set in M_{22} with the given inner product.

Definition: Orthonormal Set

A set W is said to be orthonormal if it is an orthogonal set and each vector has norm 1.

EXAMPLE 4.48 Show that the set {[1, 0, 0], [0, 1, 0], [0, 0, 1]} is an orthonormal set in R^3 with Euclidean inner product .

Solution: Let $E = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\} \subset R^3$.

To prove that E is an orthonormal set, we have to show that E is an orthogonal set and each vector has norm 1 in Euclidean inner product R^3.

(i) Let $\mathbf{e}_1 = [1, 0, 0], \mathbf{e}_2 = [0, 1, 0], \mathbf{e}_3 = [0, 0, 1] \in R^3$.

In the Euclidean inner product space R^3,

$$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = (1)(0) + (0)(1) + (0)(0) = 0$$

$$\langle \mathbf{e}_2, \mathbf{e}_3 \rangle = (0)\,(0) + (1)\,(0) + (0)\,(1) = 0$$

$$\langle \mathbf{e}_3, \mathbf{e}_1 \rangle = (0)\,(1) + (0)\,(0) + (1)\,(0) = 0$$

Hence, the set E is an orthogonal set.

(ii) In the Euclidean inner product space R^3,

$$\|\mathbf{e}_1\| = \sqrt{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} = 1^2 + 0^2 + 0^2 = 1$$

$$\|\mathbf{e}_2\| = \sqrt{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} = 0^2 + 1^2 + 0^2 = 1$$

$$\|\mathbf{e}_3\| = \sqrt{\langle \mathbf{e}_3, \mathbf{e}_3 \rangle} = 0^2 + 0^2 + 1^2 = 1$$

$$\|\mathbf{e}_i\| = 1, \text{ for each } i \qquad (i = 1, 2, 3)$$

Hence E is an orthonormal set in R^3.

EXAMPLE 4.49 Show that the set $E = \left\{\left[\frac{1}{3}, 0, 0\right], \left[0, \frac{1}{2}, 0\right], [0, 0, 1]\right\}$ is an orthogonal set in R^3 with weighted inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$.

Solution: We will prove that the set E is an orthonormal set in the following two steps:

(i) First, we show that E is an orthogonal set.

Let $\mathbf{e}_1 = \left[\frac{1}{3}, 0, 0\right]$, $\mathbf{e}_2 = \left[0, \frac{1}{2}, 0\right]$, $\mathbf{e}_3 = [0, 0, 1]$

In the weighted inner product,

$$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = 9\left(\frac{1}{3}\right)(0) + 4(0)\left(\frac{1}{2}\right) + (0)\,(0) = 0$$

$$\langle \mathbf{e}_2, \mathbf{e}_3 \rangle = 9(0)\,(0) + 4\left(\frac{1}{2}\right)(0) + (0)\,(1) = 0$$

$$\langle \mathbf{e}_3, \mathbf{e}_1 \rangle = 9(0)\left(\frac{1}{3}\right) + 4(0)\,(0) + (1)\,(0) = 0$$

Thus E is an orthogonal set.

(ii) $$\|\mathbf{e}_1\| = \sqrt{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} = \sqrt{9\left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + 4(0)\,(0) + (0)\,(0)} = 1$$

$$\|\mathbf{e}_2\| = \sqrt{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} = \sqrt{9(0)\,(0) + 4\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + (0)\,(0)} = 1$$

$$\|\mathbf{e}_3\| = \sqrt{\langle \mathbf{e}_3, \mathbf{e}_3 \rangle} = \sqrt{9(0)\,(0) + 4(0)\,(0) + (1)\,(1)} = 1$$

That is, $\|\mathbf{e}_i\| = 1$, for each $i = 1, 2, 3$

Hence E is an orthonormal set in R^3 with the given weighted Euclidean product.

$$\langle \mathbf{E}_2, \mathbf{E}_3 \rangle = (0)(0) + (1)(0) + (0)(1) + (0)(0) = 0$$

$$\langle \mathbf{E}_2, \mathbf{E}_4 \rangle = (0)(0) + (1)(0) + (0)(0) + (0)(1) = 0$$

$$\langle \mathbf{E}_3, \mathbf{E}_4 \rangle = (0)(0) + (0)(0) + (1)(0) + (0)(1) = 0$$

$$\langle \mathbf{E}_i, \mathbf{E}_j \rangle = 0, \qquad \text{for all } i \neq j,\ i, j = 1, 2, 3, 4$$

Thus E is an orthogonal set.

(ii) Unit norms:

$$\|\mathbf{E}_1\| = \sqrt{\langle \mathbf{E}_1, \mathbf{E}_1 \rangle} = \sqrt{1^2 + 0^2 + 0^2 + 0^2} = 1$$

$$\|\mathbf{E}_2\| = \sqrt{\langle \mathbf{E}_2, \mathbf{E}_2 \rangle} = \sqrt{0^2 + 1^2 + 0^2 + 0^2} = 1$$

$$\|\mathbf{E}_3\| = \sqrt{\langle \mathbf{E}_3, \mathbf{E}_3 \rangle} = \sqrt{0^2 + 0^2 + 1^2 + 0^2} = 1$$

$$\|\mathbf{E}_4\| = \sqrt{\langle \mathbf{E}_4, \mathbf{E}_4 \rangle} = \sqrt{0^2 + 0^2 + 0^2 + 1^2} = 1$$

That is, $\|\mathbf{E}_i\| = 1, \qquad$ for each $i = 1, 2, 3, 4$

Hence E is an orthonormal set in M_{22}.

Recall that any nonzero vector $\mathbf{u}$ in the inner product space can be normalized by dividing it by $\|\mathbf{u}\|$, i.e. we can find unit vector from the nonzero vector by the process of a vector dividing by its norm. In other words, if $\mathbf{u}$ is a nonzero vector in the inner product space V, then $\dfrac{\mathbf{u}}{\|\mathbf{u}\|}$ is a unit vector in V.

$$\left\| \frac{\mathbf{u}}{\|\mathbf{u}\|} \right\| = \frac{\|\mathbf{u}\|}{\|\mathbf{u}\|} = 1$$

Therefore, every orthogonal set can always be converted to an orthonormal set by normalizing each vector of it.

EXAMPLE 4.52 Convert the orthogonal set given in Example 4.43 to an orthonormal set in Euclidean inner product space R^3.

Solution: Since $W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, where $\mathbf{w}_1 = [3, 1, 1]$, $\mathbf{w}_2 = [-1, 2, 1]$, $\mathbf{w}_3 = \left[-\frac{1}{2}, -2, \frac{7}{2}\right]$

To convert the set W to an orthonormal set, we will normalize each vector $\mathbf{w}_i$ of W.

$$\|\mathbf{w}_1\| = \sqrt{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} = \sqrt{3^2 + 1^2 + 1^2} = \sqrt{11}$$

$$\|\mathbf{w}_2\| = \sqrt{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} = \sqrt{(-1)^2 + 2^2 + 1^2} = \sqrt{6}$$

$$\|\mathbf{w}_3\| = \sqrt{\langle \mathbf{w}_3, \mathbf{w}_3 \rangle} = \sqrt{\left(-\frac{1}{2}\right)^2 + (-2)^2 + \left(\frac{7}{2}\right)^2} = \sqrt{\frac{1}{4} + 4 + \frac{49}{4}} = \sqrt{\frac{33}{2}}$$

Normalized vectors:

$$\mathbf{w}_1' = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \left[\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right]$$

$$\mathbf{w}_2' = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left[-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right]$$

$$\mathbf{w}_3' = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \left[-\frac{1}{2}\sqrt{\frac{2}{33}}, -2\sqrt{\frac{2}{33}}, \frac{7}{2}\sqrt{\frac{2}{33}}\right]$$

$$= \left(-\frac{1}{\sqrt{66}}, -\frac{4}{\sqrt{66}}, \frac{7}{\sqrt{66}}\right)$$

Therefore the converted orthonormal set is $W' = \{\mathbf{w}_1', \mathbf{w}_2', \mathbf{w}_3'\}$.

EXAMPLE 4.53 Convert the orthogonal set given in Example 4.44 to an orthonormal set in weighted Euclidean inner product R^3.

Solution: Since $W = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ where $\mathbf{w}_1 = [1, 0, -1]$, $\mathbf{w}_2 = [1, 0, 9]$, $\mathbf{w}_3 = [0, 1, 0]$. Also the weighted inner product in R^3 is given by the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$$

Now we calculate the norms of vectors of W with respect to the given inner product.

$$\|\mathbf{w}_1\| = \sqrt{9(1)^2 + 4(0)^2 + (-1)^2} = \sqrt{10}$$

$$\|\mathbf{w}_2\| = \sqrt{9(1)^2 + 4(0)^2 + (9)^2} = \sqrt{90}$$

$$\|\mathbf{w}_3\| = \sqrt{9(0)^2 + 4(1)^2 + (0)^2} = 2$$

Normalized vectors:

$$\mathbf{w}_1' = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \left(\frac{1}{\sqrt{10}}, 0, -\frac{1}{\sqrt{10}}\right)$$

$$\mathbf{w}_2' = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left(\frac{1}{\sqrt{90}}, 0, -\frac{9}{\sqrt{90}}\right) = \left(\frac{1}{3\sqrt{10}}, 0, \frac{-3}{\sqrt{10}}\right)$$

$$\mathbf{w}_3' = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \left(0, \frac{1}{2}, 0\right)$$

Therefore the converted orthonormal set is $W' = \{\mathbf{w}_1', \mathbf{w}_2', \mathbf{w}_3'\}$ in R^3 with a weighted inner product.

EXAMPLE 4.54 Find the orthonormal set from the orthogonal set given in Example 4.45 in P_2 with an inner product $\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2$.

Solution: Here $P = \{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$ where $\mathbf{p}(x) = \frac{4}{5}x + \frac{3}{5}x^2$, $\mathbf{q}(x) = -\frac{3}{5}x + \frac{4}{5}x^2$, $\mathbf{r}(x) = 1$

To find the orthonormal set from the orthogonal set P, we have to normalize each vector of the set P. First, we find the norm of each vector with respect to the given inner product.

$$\|\mathbf{p}\| = \sqrt{\langle \mathbf{p}, \mathbf{p} \rangle} = \sqrt{0^2 + \left(\frac{4}{5}\right)^2 + \left(\frac{3}{5}\right)^2} = 1$$

$$\|\mathbf{q}\| = \sqrt{\langle \mathbf{q}, \mathbf{q} \rangle} = \sqrt{0^2 + \left(-\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = 1$$

$$\|\mathbf{r}\| = \sqrt{\langle \mathbf{r}, \mathbf{r} \rangle} = \sqrt{1^2 + 0^2 + 0^2} = 1$$

Normalized vectors:

$$\mathbf{p}' = \frac{\mathbf{p}}{\|\mathbf{p}\|} = \mathbf{p} = \frac{4}{5}x + \frac{3}{5}x^2$$

$$\mathbf{q}' = \frac{\mathbf{q}}{\|\mathbf{q}\|} = \mathbf{q} = -\frac{3}{5}x + \frac{4}{5}x^2$$

$$\mathbf{r}' = \frac{\mathbf{r}}{\|\mathbf{r}\|} = \mathbf{r} = 1$$

Therefore the set $P = \{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$ is itself an orthonormal set in P_2.

EXAMPLE 4.55 Find the orthonormal set from the orthogonal set given in Example 4.47 in M_{22}.

Solution: Here $S = \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$

where $\mathbf{V}_1 = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$, $\mathbf{V}_2 = \begin{bmatrix} -2 & 2 \\ 3 & 2 \end{bmatrix}$, $\mathbf{V}_3 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, $\mathbf{V}_4 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$.

Also an inner product in M_{22} is given by the formula

$$\langle \mathbf{U}, \mathbf{V} \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

(i) Norms of the vectors of set *S:*

$$\|\mathbf{V}_1\| = \sqrt{\langle \mathbf{V}_1, \mathbf{V}_1 \rangle} = \sqrt{1^2 + (-1)^2 + 2^2 + (-1)^2} = \sqrt{7}$$

$$\|\mathbf{V}_2\| = \sqrt{\langle \mathbf{V}_2, \mathbf{V}_2 \rangle} = \sqrt{(-2)^2 + 2^2 + 3^2 + 2^2} = \sqrt{21}$$

$$\|\mathbf{V}_3\| = \sqrt{\langle \mathbf{V}_3, \mathbf{V}_3 \rangle} = \sqrt{1^2 + 2^2 + 0^2 + (-1)^2} = \sqrt{6}$$

$$\|\mathbf{V}_4\| = \sqrt{\langle \mathbf{V}_4, \mathbf{V}_4 \rangle} = \sqrt{1^2 + 0^2 + 0^2 + 1^2} = \sqrt{2}$$

(ii) Normalized vectors of set S:

$$\mathbf{V}_1' = \frac{\mathbf{V}_1}{\|\mathbf{V}_1\|} = \frac{1}{\sqrt{7}}\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \\ \frac{2}{\sqrt{7}} & -\frac{1}{\sqrt{7}} \end{bmatrix}$$

$$\mathbf{V}_2' = \frac{\mathbf{V}_2}{\|\mathbf{V}_2\|} = \frac{1}{\sqrt{21}}\begin{bmatrix} -2 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{21}} & \frac{2}{\sqrt{21}} \\ \frac{3}{\sqrt{21}} & \frac{2}{\sqrt{21}} \end{bmatrix}$$

$$\mathbf{V}_3' = \frac{\mathbf{V}_3}{\|\mathbf{V}_3\|} = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{6}} \end{bmatrix}$$

$$\mathbf{V}_4' = \frac{\mathbf{V}_4}{\|\mathbf{V}_4\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Therefore $S' = \{\mathbf{V}_1', \mathbf{V}_2', \mathbf{V}_3', \mathbf{V}_4'\}$ is an orthonormal set in M_{22}.

Theorem 4.6

If $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is an orthogonal set of nonzero vectors in an inner product space V, then W is a linearly independent set in V.

Corollary 4.1 If $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is an orthonormal set of nonzero vectors in an inner product space V, then W is a linearly independent set in V.

Definition: Orthogonal Basis

An orthogonal set of an inner product space which is also a basis is called an orthogonal basis for an inner product space.

Definition: Orthonormal Basis

An orthonormal set of an inner product space which is also a basis is called an orthonormal basis for an inner product space.

Recall that a linearly independent set with n-vectors of an n-dimensional vector space is a basis for the vector space.

From Theorem 4.6, an orthogonal set with n vectors is a linearly independent set of n vectors in an n-dimensional inner product space . Therefore, an orthogonal set with n vectors in an n-dimensional inner product space is an orthogonal basis for it.

Theorem 4.7

(i) If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is an orthonormal basis for an inner product space V and $\mathbf{w}$ is any vector in V, then

$$\mathbf{w} = \frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle \mathbf{v}_1}{\|\mathbf{v}_1\|^2} + \frac{\langle \mathbf{w}, \mathbf{v}_2 \rangle \mathbf{v}_2}{\|\mathbf{v}_2\|^2} + \cdots + \frac{\langle \mathbf{w}, \mathbf{v}_n \rangle \mathbf{v}_n}{\|\mathbf{v}_n\|^2}$$

where $\frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle \mathbf{v}_1}{\|\mathbf{v}_1\|^2}, \frac{\langle \mathbf{w}, \mathbf{v}_2 \rangle \mathbf{v}_2}{\|\mathbf{v}_2\|^2}, \ldots, \frac{\langle \mathbf{w}, \mathbf{v}_n \rangle \mathbf{v}_n}{\|\mathbf{v}_n\|^2}$ are the co-ordinates of the vector $\mathbf{w}$ relative to the orthogonal basis B. that is, $[\mathbf{w}]_B = \left[\frac{\langle \mathbf{w}, \mathbf{v}_1 \rangle \mathbf{v}_1}{\|\mathbf{v}_1\|^2}, \frac{\langle \mathbf{w}, \mathbf{v}_2 \rangle \mathbf{v}_2}{\|\mathbf{v}_2\|^2}, \ldots, \frac{\langle \mathbf{w}, \mathbf{v}_n \rangle \mathbf{v}_n}{\|\mathbf{v}_n\|^2}\right]$

(ii) If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V and $\mathbf{w}$ is any vector in V then $\mathbf{w} = \langle \mathbf{w}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{w}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{w}, \mathbf{v}_n \rangle \mathbf{v}_n$ where $\langle \mathbf{w}, \mathbf{v}_1 \rangle, \langle \mathbf{w}, \mathbf{v}_2 \rangle, \cdots,$ $\langle \mathbf{w}, \mathbf{v}_n \rangle$ are the co-ordinates of the vector $\mathbf{w}$ relative to the orthonormal basis B. That is, $[\mathbf{w}]_B = \{\langle \mathbf{w}, \mathbf{v}_1 \rangle, \langle \mathbf{w}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{w}, \mathbf{v}_n \rangle\}$

EXAMPLE 4.56 Show that a set $E = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ is an orthonormal basis in R^3. Also find the co-ordinates of vector [2, 4, 3] relative to E.

Solution: In Example 4.48, we proved that E is an orthonormal set and also, it contains three vectors in the 3-dimensional space R^3. Therefore, E is an orthonormal basis in R^3.

Let us consider $\mathbf{e}_1 = [1, 0, 0]$, $\mathbf{e}_2 = [0, 1, 0]$, $\mathbf{e}_3 = [0, 0, 1]$ and $\mathbf{w} = [2, 4, 3]$.

The co-ordinates of $\mathbf{w}$ relative to orthonormal basis E are

$$\langle \mathbf{w}, \mathbf{e}_1 \rangle = (2)\,(1) + (4)\,(0) + (3)\,(0) = 2$$

$$\langle \mathbf{w}, \mathbf{e}_2 \rangle = (2)\,(0) + (4)\,(1) + (3)\,(0) = 4$$

$$\langle \mathbf{w}, \mathbf{e}_3 \rangle = (2)\,(0) + (4)\,(0) + (3)\,(1) = 3$$

That is $[\mathbf{u}]_E = [2, 4, 3] = 2\mathbf{e}_1 + 4\mathbf{e}_2 + 3\mathbf{e}_3$. Note that the set $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is called the *standard orthonormal* basis for R^3 with Euclidean inner product.

EXAMPLE 4.57 Show that set $W' = \left\{\left[\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right], \left[-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right], \left[-\frac{1}{\sqrt{66}}, -\frac{4}{\sqrt{66}}, \frac{7}{\sqrt{66}}\right]\right\}$ is an orthonormal basis for Euclidean inner product space R^3. Find the co-ordinates of [2, 4, 3] relative to W'.

Solution: In Example 4.52, we showed that a set W is an orthogonal set and also it contains three vectors in 3-dimensional space R^3. Therefore, the set W is an orthonormal basis for R^3.

Let $\mathbf{w}_1' = \left[\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right]$, $\mathbf{w}_2' = \left[-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right]$, $\mathbf{w}_3' = \left[-\frac{1}{\sqrt{66}}, -\frac{4}{\sqrt{66}}, \frac{7}{\sqrt{66}}\right]$ and $\mathbf{u} = [2, 4, 3]$.

The co-ordinates of **u** relative to an orthogonal basis W' are

$$\langle \mathbf{u}, \mathbf{w}_1' \rangle = \frac{(2)(3)+(4)(1)+(3)(1)}{\sqrt{11}} = \frac{13}{\sqrt{11}}$$

$$\langle \mathbf{u}, \mathbf{w}_2' \rangle = \frac{(2)(-1)+(4)(2)+(3)(1)}{\sqrt{6}} = \frac{9}{\sqrt{6}}$$

$$\langle \mathbf{u}, \mathbf{w}_3' \rangle = \frac{(2)(-1)+(4)(-4)+(3)(7)}{\sqrt{66}} = \frac{3}{\sqrt{66}}$$

i.e. $$\mathbf{u} = \frac{13}{\sqrt{11}}\mathbf{w}_1' + \frac{9}{\sqrt{6}}\mathbf{w}_2' + \frac{3}{\sqrt{66}}\mathbf{w}_3'$$

or $$[\mathbf{u}]_{W'} = \left[\frac{13}{\sqrt{11}}, \frac{9}{\sqrt{6}}, \frac{3}{\sqrt{66}}\right]$$

EXAMPLE 4.58 Prove that a set $E = \left\{\left[\frac{1}{3}, 0, 0\right], \left[0, \frac{1}{2}, 0\right], [0, 0, 1]\right\}$ is an orthonormal basis in R^3 with weighted inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$. Also find the co-ordinates of the vector [1, –2, 3].

Solution: We have seen in Example 4.49 that E is an orthonormal set and also it has three vectors. So, E is an orthonormal basis for R^3. Let $\mathbf{w}_1 = \left[\frac{1}{3}, 0, 0\right], \mathbf{w}_2 = \left[0, \frac{1}{2}, 0\right], \mathbf{w}_3 = [0, 0, 1]$ relative to the orthonormal basis E.

$$\langle \mathbf{u}, \mathbf{w}_1 \rangle = 9(1)\left(\frac{1}{3}\right) + 4(-2)(0) + (3)(0) = 3$$

$$\langle \mathbf{u}, \mathbf{w}_2 \rangle = 9(1)(0) + 4(-2)\left(\frac{1}{2}\right) + (3)(0) = -4$$

$$\langle \mathbf{u}, \mathbf{w}_3 \rangle = 9(1)(0) + 4(-2)(0) + (3)(1) = 3$$

Therefore, $\mathbf{u} = 3\mathbf{w}_1 - 4\mathbf{w}_2 + 3\mathbf{w}_3$ i.e. $[\mathbf{u}]_E = [3, -4, 3]$.

EXAMPLE 4.59 Is the set $E = \{1, x, x^2\}$ an orthonormal basis for P_2 with an inner product $\langle \mathbf{p}, \mathbf{q} \rangle = p_0q_0 + p_1q_1 + p_2q_2$? Find the co-ordinates of $5 - 2x + 3x^2$.

Solution: From Example 4.50, E is an orthonormal set and it has three-vectors in a 3-dimensional space P_2. Therefore, E is an orthonormal basis for P_2.

The co-ordinates of the vector $\mathbf{u} = 5 - 2x + 3x^2$ are

$$\langle \mathbf{u}, \mathbf{e}_1 \rangle = (5)(1) + (-2)(1) + (3)(0) = 5$$

$$\langle \mathbf{u}, \mathbf{e}_2 \rangle = (5)(0) + (-2)(1) + (3)(0) = -2$$

EXAMPLE 4.63 Verify the result (ii) of Theorem 4.8, $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$, for a vector $\mathbf{u} = [1, -2, 3]$ relative to orthonormal basis $E = \left\{\left[\frac{1}{3}, 0, 0\right], \left[0, \frac{1}{2}, 0\right], [0, 0, 1]\right\}$ in R^3 with weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$.

Solution: Let $\mathbf{u} = [1, -2, 3] \in R^3$.
The norm of $\mathbf{u}$ in R^3 with respect to the given weighted inner product is

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{9(1)^2 + 4(-2)^2 + (3)^2} = \sqrt{9+16+9} = \sqrt{34} \tag{i}$$

We have already calculated the co-ordinates of $\mathbf{u} = [1, -2, 3]$ in Example 4.58 relative to E as

$$[\mathbf{u}]_E = [u_1, u_2, u_3] = [3, -4, 3].$$

$$\sqrt{u_1^2 + u_2^2 + u_3^2} = \sqrt{9+16+9} = \sqrt{34} \tag{ii}$$

From (i) and (ii), we have

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$$

Orthogonal Projection

First we will introduce the concept of orthogonal projection of one vector on the other vector. Then we will extend the concept of orthogonal projection of a vector on the subspace of an inner product space.

To understand the concept of orthogonal projection of a vector $\mathbf{u}$ on the vector $\mathbf{w}$ of an inner product space V, consider the problem of decomposing a vector $\mathbf{u}$ into the sum of two vectors, one a multiple of $\mathbf{w}$ and the other orthogonal to $\mathbf{w}$ (Figure 4.1), that is,

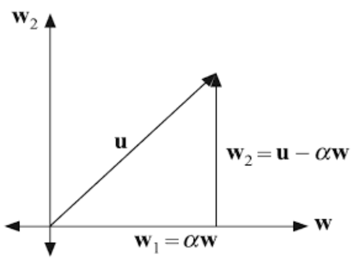


Figure 4.1 Orthogonal projection.

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{4.1}$$

where $\mathbf{w}_1$ is a multiple of $\mathbf{w}$, $\mathbf{w}_1 = \alpha\mathbf{w}$ for some scalar α and $\mathbf{w}_2$ is some vector orthogonal to $\mathbf{w}$.

Here $\mathbf{w}_1 = \alpha\mathbf{w}$

$\mathbf{w}_2 = \mathbf{u} - \alpha\mathbf{w}$ [from Eq. (4.1)]

$\mathbf{u} - \alpha\mathbf{w}$ is orthogonal to $\mathbf{w}$ (since $\mathbf{w}_2$ is orthogonal to $\mathbf{w}$)

that is, $\langle \mathbf{w}, \mathbf{u} - \alpha\mathbf{w} \rangle = 0$

$$\langle \mathbf{w}, \mathbf{u} \rangle - \alpha\langle \mathbf{w}, \mathbf{w} \rangle = 0$$

$$\alpha = \frac{\langle \mathbf{w}, \mathbf{u} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} = \frac{\langle \mathbf{w}, \mathbf{u} \rangle}{\|\mathbf{w}\|^2}$$

$$\mathbf{w}_1 = \frac{\langle \mathbf{w}, \mathbf{u} \rangle}{\|\mathbf{w}\|^2}\mathbf{w} \quad \text{and} \quad \mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1 = \mathbf{u} - \frac{\langle \mathbf{w}, \mathbf{u} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}$$

The vector $\mathbf{w}_1$ is called the orthogonal projection of a vector $\mathbf{u}$ on the vector $\mathbf{w}$ and $\mathbf{w}_2$ is called the component of $\mathbf{u}$ orthogonal to $\mathbf{w}$.

Solution: It is easy to check that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for the subspace W of R^3.

Here $\mathbf{u} = [5, 2, 10] \in R^3$. From Theorem 4.6,

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \quad \text{where } \mathbf{w}_1 = \text{Proj}_W \mathbf{u} \in W \text{ and } \mathbf{w}_2 = \mathbf{u} - \text{Proj}_W \mathbf{u} \in W^{\perp}$$

Since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for W. Therefore, by Theorem 4.4

$$\begin{aligned}
\mathbf{w}_1 &= \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\
&= \frac{(5)(2) + (2)(5) + (10)(-1)}{(4 + 25 + 1)}[2, 5, -1] + \frac{(5)(-2) + (2)(1) + (10)(1)}{(4 + 1 + 1)}[-2, 1, 1] \\
&= \frac{10}{30}[2, 5, -1] + \frac{2}{6}[-2, 1, 1] \\
&= \frac{1}{3}[2, 5, -1] + \frac{1}{3}[-2, 1, 1] \\
&= [0, 2, 0]
\end{aligned}$$

$$\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1 = [5, 2, 10] - [0, 2, 0] = [5, 0, 10]$$

Therefore $\quad \mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 = [0, 2, 0] + [5, 0, 10] = [5, 2, 10]$

where $\mathbf{w}_1 = [0, 2, 0] \in W$ and $\mathbf{w}_2 = [5, 0, 10] \in W^{\perp}$.

EXAMPLE 4.66 If W is a subspace of a Euclidean space R^4 spanned by the vectors $\mathbf{v}_1 = [1, 1, 0, -1]$, $\mathbf{v}_2 = [1, 0, 1, 1]$, $\mathbf{v}_3 = [0, -1, 1, -1]$, then express $\mathbf{u} = [3, 4, 5, 6]$ as the sum of a vector in W and a vector orthogonal to W.

Solution: It is easy to show that a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for a subspace W of Euclidean space R^4. Here $\mathbf{u} = [3, 4, 5, 6] \in R^4$

From Theorem 4.6,

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$$

where $\mathbf{w}_1 = \text{proj}_W \mathbf{u} \in W$ and $\mathbf{w}_2 = \mathbf{u} - \text{proj}_W \mathbf{u} \in W^{\perp}$.

Since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W, we have by Theorem 4.4

$$\begin{aligned}
\mathbf{w}_1 &= \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \frac{\langle \mathbf{u}, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3 \\
&= \frac{(3)(1) + (4)(1) + (5)(0) + (6)(-1)}{(1 + 1 + 0 + 1)}[1, 1, 0, -1] + \frac{(3)(1) + (4)(0) + (5)(1) + (6)(1)}{(1 + 0 + 1 + 1)}[1, 0, 1, 1] \\
&\quad \frac{(3)(0) + (4)(-1) + (5)(1) + (6)(-1)}{(0 + 1 + 1 + 1)}[0, -1, 1, -1] \\
&= \frac{1}{3}[1, 1, 0, -1] + \frac{14}{3}[1, 0, 1, 1] - \frac{5}{3}[0, -1, 1, -1]
\end{aligned}$$

$$= \frac{1}{3}[1 + 14 + 0, 1 + 0 + 5, 0 + 14 - 5, -1 + 14 + 5]$$

$$= \frac{1}{3}[15, 6, 9, 18]$$

$$= [5, 2, 3, 6]$$

Now $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1 = [3, 4, 5, 6] - [5, 2, 3, 6] = [-2, 2, 2, 0]$

Therefore $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 = [5, 2, 3, 6] + [-2, 2, 2, 0]$ where $\mathbf{w}_1 \in W$ and $\mathbf{w}_2 \in W^{\perp}$.

The next theorem is very important for the finite-dimensional inner product spaces.

Theorem 4.12

Every finite dimensional inner product space has an orthonormal basis

The proof of the above theorem is a stepwise process which constructs the orthogonal basis from an arbitrary basis of a finite dimensional inner product space. This process is called Gram–Schmidt process. The steps of this process are given below which will be used to solve some problems.

Gram–Schmidt Process

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an arbitrary basis for an inner product space V. Here the problem is to convert this arbitrary basis into an orthonormal basis for V.

The steps of Gram–Schmidt process to construct this orthonormal basis are given below.

Step 1 Take $\mathbf{w}_1 = \mathbf{v}_1$

Step 2 Compute $\mathbf{w}_2 = \mathbf{v}_2 - \dfrac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1$ which is orthogonal to $\mathbf{w}_1$.

Step 3 Compute $\mathbf{w}_3 = \mathbf{v}_3 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2\}$.

Step 4 Compute $\mathbf{w}_4 = \mathbf{v}_4 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_3 \rangle}{\|\mathbf{w}_3\|^2}\mathbf{w}_3$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

… … … … …

Step n Compute $\mathbf{w}_n = \mathbf{v}_n - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \cdots - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\|\mathbf{w}_{n-1}\|^2}\mathbf{w}_{n-1}$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots, \mathbf{w}_{n-1}\}$.

Step (n + 1) The set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots, \mathbf{w}_n\}$ is an orthogonal basis for V.

Step (n + 2) The set $\left\{\dfrac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \dfrac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \ldots, \dfrac{\mathbf{w}_n}{\|\mathbf{w}_n\|}\right\}$ is an orthonormal basis for V.

EXAMPLE 4.67 Find the orthonormal basis from a basis {[1, 1, 1], [1, 2, 1], [–1, 1, 0]} for a Euclidean inner product space R^3.

Solution: Let $\mathbf{v}_1 = [1, 1, 1]$, $\mathbf{v}_2 = [1, 2, 1]$, $\mathbf{v}_3 = [-1, 1, 0] \in R^3$.

A set $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for R^3.

We want to convert B into an orthonormal basis for R^3 by using the Gram–Schmidt process.

Step 1 Take $\mathbf{w}_1 = \mathbf{v}_1 = [1, 1, 1]$

Step 2
$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1$$
$$= [1, 2, 1] - \frac{(1)(1) + (2)(1) + (1)(1)}{(1+1+1)}[1, 1, 1]$$
$$= [1, 2, 1] - \left[\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right] = \left[-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right]$$

Step 3
$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2$$
$$= [-1, 1, 0] - \frac{-1+1+0}{(1+1+1)}[1, 1, 1] - \frac{\frac{1}{3} + \frac{2}{3} + 0}{\left(\frac{1}{9} + \frac{4}{9} + \frac{1}{9}\right)}\left[-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right]$$
$$= [-1, 1, 0] - 0 - \frac{3}{2}\left[-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right]$$
$$= [-1, 1, 0] - \left[-\frac{1}{2}, 1, -\frac{1}{2}\right]$$
$$= \left[-\frac{1}{2}, 0, \frac{1}{2}\right]$$

Step 4 $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal basis for R^3

Step 5
$$\left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}\right\} = \left\{\frac{1}{\sqrt{3}}[1, 1, 1], \frac{3}{\sqrt{6}}\left[-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right], \sqrt{2}\left[-\frac{1}{2}, 0, \frac{1}{2}\right]\right\}$$
$$= \left\{\left[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right], \left[-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right], \left[-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right]\right\}$$

It is an orthonormal basis for R^3.

EXAMPLE 4.68 Construct an orthogonal basis of the subspace spanned by the vectors $\mathbf{v}_1 = [1, -4, 0, 1]$, $\mathbf{v}_2 = [7, -7, -4, 1]$ of a Euclidean inner product space R^4.

$$= \left(\frac{x^3}{3}\right)_{-1}^{1}$$

$$= \frac{2}{3}$$

Step 3 $\quad \mathbf{w}_3 = \mathbf{v}_3 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2$

$$= x^2 - \frac{1}{2}\left(\int_{-1}^{1} v_3(x)w_1(x)dx\right)\mathbf{w}_1 - \frac{3}{2}\left(\int_{-1}^{1} v_3(x)w_2(x)dx\right)\mathbf{w}_2$$

$$= x^2 - \frac{1}{2}\left(\int_{-1}^{1} x^2 dx\right)(1) - \frac{3}{2}\left(\int_{-1}^{1} x^3 dx\right)x$$

$$= x^2 - \frac{1}{2}\left(\frac{x^3}{3}\right)_{-1}^{1} - \frac{3}{2}(0)$$

$$= x^2 - \frac{1}{2}\left(\frac{2}{3}\right)$$

$$= x^2 - \frac{1}{3}$$

$$\therefore \qquad \mathbf{w}_3 = x^2 - \frac{1}{3}$$

Now $\|\mathbf{w}_3\|^2 = \langle \mathbf{w}_3, \mathbf{w}_3 \rangle$

$$= \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx$$

$$= \left(\frac{x^5}{5} - \frac{2}{9}x^3 + \frac{1}{9}x\right)_{-1}^{1}$$

$$= \frac{2}{5} - \frac{4}{9} + \frac{2}{9}$$

$$= \frac{8}{45}$$

Step 4 $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal basis for P_2.

Step 5 The orthonormal basis for P_2 is

$$\left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}\right\} = \left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{3\sqrt{5}}{2\sqrt{2}}\left(x^2 - \frac{1}{3}\right)\right\}.$$

EXAMPLE 4.70 Consider an inner product space R^3 with an inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$. Find the orthogonal basis for R^3 from a basis {[1, 0, 1], [0, 1, 2], [2, 1, 0]}.

Solution: Let $\mathbf{v}_1 = [1, 0, 1]$, $\mathbf{v}_2 = [0, 1, 2]$, $\mathbf{v}_3 = [2, 1, 0]$

The inner product is defined by the formula $\langle \mathbf{u}, \mathbf{v} \rangle = 9u_1v_1 + 4u_2v_2 + u_3v_3$.

By using the Gram–Schmidt method:

Step 1 Take $\mathbf{w}_1 = \mathbf{v}_1 = [1, 0, 1] \Rightarrow \|\mathbf{w}_1\|^2 = 9(1)^2 + 4(0)^2 + (1)^2 = 10$

Step 2

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1$$

$$= [0, 1, 2] - \frac{9(0) + 4(0) + 2}{10}[1, 0, 1]$$

$$= [0, 1, 2] - \left[\frac{1}{5}, 0, \frac{1}{5}\right]$$

$$= \left[-\frac{1}{5}, 1, \frac{9}{5}\right]$$

$$\therefore \quad \mathbf{w}_2 = \left[-\frac{1}{5}, 1, \frac{9}{5}\right]$$

Now

$$\|\mathbf{w}_2\|^2 = 9\left(\frac{1}{5}\right)^2 + 4(1)^2 + \left(\frac{9}{5}\right)^2$$

$$= \frac{9}{25} + 4 + \frac{81}{25}$$

$$= \frac{38}{5}$$

Step 3

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_2$$

$$= [2, 1, 0] - \frac{18}{10}[1, 0, 1] - \frac{2}{38}\left[-\frac{1}{5}, 1, \frac{9}{5}\right]$$

$$= [2, 1, 0] - \left[\frac{9}{5}, 0, \frac{9}{5}\right] - \left[-\frac{1}{95}, \frac{1}{19}, \frac{9}{95}\right]$$

$$= \left[2 - \frac{9}{5} + \frac{1}{95}, 1 - \frac{1}{19}, -\frac{9}{5} - \frac{9}{95}\right]$$

$$= \left[\frac{4}{19}, \frac{18}{19}, -\frac{36}{19}\right]$$

Step 4 $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal basis for R^3 with the given inner product.

Therefore the matrix $\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

For the upper triangular matrix,

$$\mathbf{R} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{q}_1 \rangle & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle & \langle \mathbf{v}_3, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2 \rangle & \langle \mathbf{v}_3, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{v}_3, \mathbf{q}_3 \rangle \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{3} & \frac{4}{\sqrt{3}} & 0 \\ 0 & \frac{2}{\sqrt{6}} & \frac{3}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Hence the **QR**-decomposition of matrix **A** is $\mathbf{A} = \mathbf{QR} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & \frac{4}{\sqrt{3}} & 0 \\ 0 & \frac{2}{\sqrt{6}} & \frac{3}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$

EXAMPLE 4.72 Find the **QR**-decomposition of $\mathbf{A} = \begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$.

Solution: The column vectors of **A** are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 4 \\ -4 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ -4 \\ -3 \\ 7 \\ 1 \end{bmatrix},$$

Now, apply the Gram–Schmidt process to find the orthogonal basis for the column space of matrix **A**.

Step 1 Take $\mathbf{w}_1 = \mathbf{v}_1 \Rightarrow \|\mathbf{w}_1\|^2 = 1 + 1 + 1 + 1 + 1 = 5$

Step 2

$$\begin{aligned}
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\
&= [2, 1, 4, -4, 2] - \frac{2-1-4-4+2}{5}[1, -1, -1, 1, 1] \\
&= [2, 1, 4, -4, 2] + [1, -1, -1, 1, 1] \\
&= [3, 0, 3, -3, 3]
\end{aligned}$$

$$\therefore \qquad \mathbf{w}_2 = [3, 0, 3, -3, 3]$$

and $\|\mathbf{w}_2\|^2 = 9 + 0 + 9 + 9 + 9 = 36$

Step 3

$$\begin{aligned}
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\
&= [5, -4, -3, 7, 1] - \frac{15+0-9-21+3}{36}[3, 0, 3, -3, 3] - \frac{5+4+3+7+1}{5}[1, -1, -1, 1, 1] \\
&= [5, -4, -3, 7, 1] + \frac{1}{3}[3, 0, 3, -3, 3] - 4[1, -1, -1, 1, 1] \\
&= [2, 0, 2, 2, -2]
\end{aligned}$$

$$\therefore \qquad \mathbf{w}_3 = [2, 0, 2, 2, -2]$$

Now $\|\mathbf{w}_3\|^2 = 4 + 0 + 4 + 4 + 4 = 16.$

Step 4 Therefore $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is an orthogonal basis of column space of **A** and $\left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}\right\}$.

$$= \left\{\left[\frac{1}{\sqrt{5}}, -\frac{1}{\sqrt{5}}, -\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right], \left[\frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right], \left[\frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right]\right\}$$

is an orthonormal basis for the column space of matrix **A**.

$$\therefore \qquad \mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{5}} & 0 & 0 \\ -\frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{5}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{5}} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

For the upper triangular matrix,

$$\mathbf{R} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{q}_1 \rangle & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle & \langle \mathbf{v}_3, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2 \rangle & \langle \mathbf{v}_3, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{v}_3, \mathbf{q}_3 \rangle \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

Hence the **QR**-decomposition of matrix **A** is

$$\mathbf{A} = \mathbf{QR} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{5}} & 0 & 0 \\ -\frac{1}{\sqrt{5}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{5}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{5}} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$$

EXAMPLE 4.73 Find the **QR**-decomposition of $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix}$.

Solution: The column vectors of **A** are

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Theorem 4.14 [The Best Approximation Theorem]

If W is a finite-dimensional subspace of an inner product space V and $\mathbf{v}$ is any fixed vector outside the subspace W, then the orthogonal projection vector $\text{proj}_W \mathbf{v}$ in the subspace W is a best approximation of $\mathbf{v}$ by the vector of subspace W (Figure 4.5).

$$\|\mathbf{v} - \text{proj}_W \mathbf{v}\| < \|\mathbf{v} - \mathbf{w}\| \qquad \text{for } \forall \mathbf{w} \in W \text{ and } \mathbf{w} \neq \text{proj}_W \mathbf{v}$$

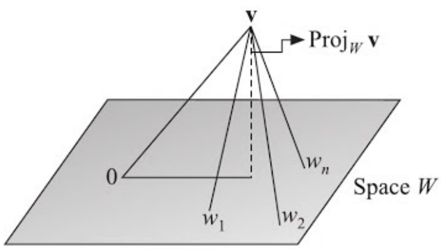


Figure 4.5 The best approximation theorem.

Definition: Error of Approximation

If the vector $\text{proj}_W \mathbf{v}$ in subspace W is called approximated *error vector*, then $\|\mathbf{v} - \text{proj}_W \mathbf{v}\|$ is called an *error of approximation*.

EXAMPLE 4.74 Find the vector on a line W: $x = 2t$, $y = t$, $z = 3t$ in R^3 which is closest to the vector [4, 5, 5]. Hence find the distance between them.

Solution: The parametric equation of a line W is: $x = 2t$, $y = t$, $z = 3t$

$$W = \{t[2, 1, 3] \mid t \in R\}$$

i.e. W is a subspace generated by a vector [2, 1, 3].

Let $\mathbf{v} = [4, 5, 5] \in R^3$. By Theorem 4.9, the closest point to vector $\mathbf{v}$ on a line W is the orthogonal projection vector $\text{proj}_W \mathbf{v}$. Since the subspace W is generated by the vector $\mathbf{u} = [2, 1, 3]$, hence

$$\begin{aligned}\text{proj}_W \mathbf{v} &= \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2}\mathbf{u} \\ &= \frac{(4)(2) + (5)(1) + (5)(3)}{4 + 1 + 9}[2, 1, 3] \\ &= [4, 2, 6]\end{aligned}$$

Hence [4, 2, 6] is the closest vector to the vector (4, 5, 5) on the given line W.

$$\begin{aligned}\text{The distance between } \mathbf{v} \text{ and } \text{Proj}_W \mathbf{v} &= \|\mathbf{v} - \text{proj}_W \mathbf{v}\| \\ &= \|[0, 3, -1]\| \\ &= \sqrt{0 + 9 + 1} \\ &= \sqrt{10}\end{aligned}$$

EXAMPLE 4.75 Approximate a vector [1, –2, 3] by the vector of a subspace $W = \{[s + t, 3s + 3t, 5s - 2t] \mid s, t \in R\}$ such that the error of approximation is minimum. Also, find the error vector.

Solution: Here the subspace

$$W = \{[s + t, 3s + 3t, 5s - 2t] \mid s, t \in R\}$$
$$= \{s[1, 3, 5] + t[1, 3, -2] \mid s, t \in R\}$$

i.e. W is a subspace generated by two vectors [1, 3, 5] and [1, 3, –2]. Also these vectors are orthogonal vectors , therefore {[1, 3, 5], [1, 3, –2]} is an orthogonal basis for the subspace W.

Let $\mathbf{v} = [1, -2, 3] \in R^3$

By Theorem 4.9, the closest vector to $\mathbf{v}$ on a subspace W is the orthogonal projection vector $\text{proj}_W \mathbf{v}$. Since the subspace W is generated by the orthogonal vectors $\mathbf{w}_1 = [1, 3, 5]$, $\mathbf{w}_2 = [1, 3, -2]$

Therefore,
$$\text{proj}_W \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2$$

$$= \frac{1-6+15}{1+9+25}[1, 3, 5] + \frac{1-6-6}{1+9+4}[1, 3, -2]$$

$$= \frac{10}{35}[1, 3, 5] - \frac{11}{14}[1, 3, -2]$$

$$= \left[\frac{10}{35} - \frac{11}{14}, \frac{30}{35} - \frac{33}{14}, \frac{50}{35} + \frac{22}{14}\right]$$

$$= \left[-\frac{1}{2}, -\frac{3}{2}, 3\right]$$

Error vector
$$= \mathbf{v} - \text{proj}_W \mathbf{v} = [1, -2, 3] - \left[-\frac{1}{2}, -\frac{3}{2}, 3\right]$$

$$= \left[\frac{3}{2}, -\frac{1}{2}, 0\right]$$

Problem 2: ***Approximate solution of an inconsistent system of linear equations***

The system of linear equations $\mathbf{Ax} = \mathbf{b}$ in R^n which has no solution is known as the *inconsistent system*, that is, we cannot find the vector $\mathbf{x}$ in R^n which satisfy the equation $\mathbf{Ax} = \mathbf{b}$.

The inconsistent system arises in some important physical applications but it demands the solutions even though its exact solution does not exist.

So what one can do, is to find an $\mathbf{x}$ that makes $\mathbf{Ax}$ as close as possible to $\mathbf{b}$, i.e. to find an $\mathbf{x}$ that makes $\|\mathbf{b} - \mathbf{Ax}\|$ as small as possible .

Sometimes these values of $\mathbf{x}$ are known as least squares solutions of an inconsistent system. It gives us the following definition.

Definition: ***Least-Squares Solution of the Inconsistent System***

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b}$ is a vector in R^n, then a least-squares solution of the inconsistent system $\mathbf{Ax} = \mathbf{b}$ is $\hat{\mathbf{x}}$ in R^n such that $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\| \le \|\mathbf{b} - \mathbf{Ax}\|$ for all $\mathbf{x} \in R^n$.

For the inconsistent system $\mathbf{Ax} = \mathbf{b}$, $\mathbf{Ax}$ is an element of column space of matrix $\mathbf{A}$ for every $\mathbf{x}$ in R^n. Therefore the problem of finding a least-squares solution of $\mathbf{Ax} = \mathbf{b}$ is the problem to approximate a vector $\overline{\mathbf{b}}$ by the vector of column space $C(\mathbf{A})$ of matrix $\mathbf{A}$.

$$\begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

$\therefore \qquad x_1 = 1, x_2 = 1$

Therefore the solution of the normal equations is $x_1 = 1$, $x_2 = 1$, i.e. $[x_1, x_2] = [1, 1]$. Hence the least-squares solution of $\mathbf{Ax} = \mathbf{b}$ is $\hat{\mathbf{x}} = [1, 1]$. Also, the orthogonal projection of $\mathbf{b}$ on the column space of $\mathbf{A}$ is

$$\mathbf{proj}_{C(\mathbf{A})}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$$

EXAMPLE 4.77 Find the least-squares solution of $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$. Hence find the orthogonal projection of $\mathbf{b}$ on to column space of $\mathbf{A}$.

Solution: Here $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$

The normal equation of the system $\mathbf{Ax} = \mathbf{b}$ is

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 14 \\ -5 \end{bmatrix}$$

or $\qquad 3x_1 = 1,\; 3x_2 = 14,\; 3x_3 = -5$

$\therefore \qquad x_1 = \dfrac{1}{3},\; x_2 = \dfrac{14}{3},\; x_3 = -\dfrac{5}{3}$

Therefore the solution of normal equations is $\hat{\mathbf{x}} = [x_1, x_2, x_3] = \left[\frac{1}{3}, \frac{14}{3}, -\frac{5}{3}\right]$. Hence the least-squares solution of $\mathbf{Ax} = \mathbf{b}$ is $\hat{\mathbf{x}} = \left[\frac{1}{3}, \frac{14}{3}, -\frac{5}{3}\right]$.

The orthogonal projection of **b** on to column space $C(\mathbf{A})$ is

$$\text{proj}_{C(\mathbf{A})}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{14}{3} \\ -\frac{5}{3} \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 6 \end{bmatrix}.$$

EXAMPLE 4.78 Does the linear system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$ have a unique solution? Find its least-squares solution.

Solution: Here $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$.

It is easy to check that the third column $C3$ is the sum of the first and second columns C_1 and C_2. i.e. $C_3 = C_1 + C_2$. Therefore, the columns of matrix **A** are not linearly independent. So, the given system does not have a unique solution. The least-square solutions of $\mathbf{Ax} = \mathbf{b}$ are the solution of the normal equations $\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix}$$

The augmented matrix is

$$\begin{bmatrix} 2 & 0 & 2 & : & 4 \\ 0 & 2 & 2 & : & 6 \\ 2 & 2 & 4 & : & 10 \end{bmatrix}$$

Applying $R_3 - (R_1 + R_2) \sim \begin{bmatrix} 2 & 0 & 2 & : & 4 \\ 0 & 2 & 2 & : & 6 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$

$$x_1 + x_3 = 2; \qquad x_2 + x_3 = 3$$

$$\therefore \qquad x_1 = 2 - x_3; \qquad x_2 = 3 - x_3$$

The solutions of the normal equation are: $\hat{\mathbf{x}} = \begin{bmatrix} 2 - x_3 \\ 3 - x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

Hence the least-squares solutions of $\mathbf{Ax} = \mathbf{b}$ are: $\hat{x} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, for any $s \in R$.

Problem 3: ***Orthogonal projection of a vector onto subspace generated by non-orthogonal basis***

If the subspace W of an inner product space is generated by an orthogonal basis then by the theorem, we can find the orthogonal projection of any vector onto the subspace W.

But if the subspace does not have an orthogonal basis, then we have the following two methods to find the orthogonal projection

(i) Using the Gram–Schmidt process to convert the given basis into orthogonal basis and then use the theorem to find the orthogonal projection.

(ii) Use the least-squares solutions $\hat{\mathbf{x}}$ of $\mathbf{Ax} = \mathbf{b}$ where the columns of $\mathbf{A}$ are the basis vectors of the subspace W and $\mathbf{b}$ is a vector which to be projected on W, i.e. $\text{proj}_W \mathbf{b} = \mathbf{A}\hat{\mathbf{x}}$

EXAMPLE 4.79 Find the orthogonal projection of [2, –1, 5] onto the subspace of R^3 whose basis is {[–1, 3, 5], [1, 2, 4]}.

Solution: Here the subspace W of R^3 is generated by the vectors [–1, 3, 5] and [1, 2, 4]. We want to find the orthogonal projection of vector [2, –1, 5] onto the subspace W.

Let $\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 3 & 2 \\ 5 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$

Now the subspace W is a column space of matrix $\mathbf{A}$. Therefore, the orthogonal projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$ can be found by using the least-squares solutions of the linear system of equations $\mathbf{Ax} = \mathbf{b}$. That is, if $\hat{\mathbf{x}}$ is a least-squares solution of $\mathbf{Ax} = \mathbf{b}$, then

$$\text{proj}_{C(\mathbf{A})}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}}$$

$$\text{proj}_W \mathbf{b} = \text{proj}_{C(\mathbf{A})}\mathbf{b}$$

The least-squares solutions of $\mathbf{Ax} = \mathbf{b}$ are the solution of normal equations $\mathbf{A}^T\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{b}$

$$\begin{bmatrix} -1 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & 2 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 35 & 25 \\ 25 & 21 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 20 \\ 20 \end{bmatrix}$$

$\therefore$
$$x_1 = -\frac{8}{11},\ x_2 = \frac{20}{11}$$

The least-squares solutions is $\hat{\mathbf{x}} = \begin{bmatrix} -\dfrac{8}{11} \\ \dfrac{20}{11} \end{bmatrix}$.

Therefore, the orthogonal projection of **b** on W is

$$\begin{aligned} \text{Proj}_W\mathbf{b} &= \text{Proj}_{C(A)}\mathbf{b} \\ &= \begin{bmatrix} -1 & 1 \\ 3 & 2 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} -\dfrac{8}{11} \\ \dfrac{20}{11} \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{28}{11} \\ \dfrac{16}{11} \\ \dfrac{40}{11} \end{bmatrix} \end{aligned}$$

EXAMPLE 4.80 Find the orthogonal projection of [3, 5, 7, –3] onto the subspace of **b** spanned by the vectors [1, 1, 1, 1], [3, 1, 1, 3] and [5, 0, 2, 3].

Solution: The orthogonal projection of a vector **b** onto the column space $C(\mathbf{A})$ of matrix

$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$ is $\text{proj}_{C(A)}\,\mathbf{b} = \mathbf{A}\hat{\mathbf{x}}$ where $\hat{\mathbf{x}}$ is the least-squares solutions of $\mathbf{Ax} = \mathbf{b}$. By Theorem 4.10,

the least-squares solutions of $\mathbf{Ax} = \mathbf{b}$ are the solution of normal equations

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 3 \\ 5 & 0 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 3 \\ 5 & 0 & 2 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 8 & 10 \\ 8 & 20 & 26 \\ 10 & 26 & 38 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ 12 \\ 20 \end{bmatrix}$$

The augmented matrix of the above system is

$$\left[\begin{array}{ccccc} 4 & 8 & 10 & : & 12 \\ 8 & 20 & 26 & : & 12 \\ 10 & 26 & 38 & : & 20 \end{array}\right] \sim \left[\begin{array}{ccccc} 1 & 0 & 0 & : & 10 \\ 0 & 1 & 0 & : & -6 \\ 0 & 0 & 1 & : & 2 \end{array}\right]$$

$\therefore$ $$x_1 = 10,\ x_2 = -6,\ x_3 = 2$$

The least-squares solution is $\hat{\mathbf{x}} = [10, -6, 2]$

Therefore, the orthogonal projection of $\mathbf{b}$ on W is

$$\begin{aligned} \text{proj}_W\, \mathbf{b} &= \text{proj}_{C(\mathbf{A})}\, \mathbf{b} \\ &= \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}\begin{bmatrix} 10 \\ -6 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 4 \\ 8 \\ -2 \end{bmatrix} \end{aligned}$$

EXERCISE SET 5

1. Use the least-squares solution to find the equation of the line that will best approximate the points $(-3, 70)$, $(1, 21)$, $(-7, 110)$ and $(5, -35)$.

   [**Hint:** Use the line equation $y = mx + c$]

2. Find the least-squares solution of the system $\begin{bmatrix} 2 & -1 & 1 \\ 1 & -5 & 2 \\ -3 & 1 & -4 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ 5 \\ -1 \end{bmatrix}$

3. Find the orthogonal projection of $[-3, -3, 8, 9]$ onto the subspace of R^4 whose basis is $\{[3, 1, 0, 1], [1, 2, 1, 1], (-1, 0, 2, -1]\}$.

SUMMARY

Inner Product A function defined from $V \times V$ into R is called inner product on a real vector space V if it assigns a unique real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors **u** and **v** in V that satisfies the following axioms for all vectors **u**, **v**, **w** in V and all real scalars α.

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ [Commutative/symmetry]

(ii) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ [Distributive/linearity]

(iii) $\langle \alpha\mathbf{u}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$ [Homogeneity]

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ [Positivity] and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Inner Product Space A real vector space with an inner product is called an inner product space.

Inner Product Generated by Matrix If **A** is an invertible matrix of order $n \times n$ and $\langle \mathbf{u}, \mathbf{u} \rangle$ is the Euclidean inner product on R^n, then the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{Au} \cdot \mathbf{Av} \qquad \text{where } \mathbf{u} \text{ and } \mathbf{v} \text{ are vectors of } R^n$$

defines an inner product, which is called the *inner product on R^n generated by* **A**.

Theorem [Properties of Inner Product Space]

Let V be an inner product space and **u**, **v** and **w** be three vectors of V. Let α be a scalar. Then

(i) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(ii) $\langle \mathbf{u}, \alpha\mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle$

(iii) $\langle \mathbf{u}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{u} \rangle = 0$

Norm The norm of a vector **u** in an inner product space V is given by the formula $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$.

Distance The distance between any two vectors **u** and **v** of an inner product space is given by the formula $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$.

Theorem 2 [Properties of Norm and Distance]

Let V be an inner product space. Then for arbitrary vectors **u**, **v** and **z** in V and scalar α

(i) $\|\alpha\mathbf{u}\| = |\alpha|\,\|\mathbf{u}\|$

(ii) $\|\mathbf{u}\| \geq 0$ and $\|\mathbf{u}\| = 0$ if $\mathbf{u} = \mathbf{0}$

(iii) Cauchy–Schwarz inequality $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$

(iv) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

(v) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

(vi) $d(\mathbf{u}, \mathbf{v}) \geq 0$ and $d(\mathbf{u}, \mathbf{v}) = 0$ if $\mathbf{u} = \mathbf{v}$

(vii) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{z}) + d(\mathbf{z}, \mathbf{v})$.

Angle If θ is an angle between the vectors **u** and **v** in inner product space V, then $\theta = \cos^{-1}\left(\dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)$.

(ii) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ are orthonormal basis for W, then

$$\text{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_r \rangle \mathbf{v}_r$$

$$= \sum_{i=1}^{r} \langle \mathbf{u}, \mathbf{v}_i \rangle \mathbf{v}_i$$

Theorem Let W be a finite-dimensional subspace of an inner product space V and $\mathbf{u}$ be any vector in V.

(i) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ are orthogonal basis for W, then the vector

$$\mathbf{w}_2 = \mathbf{u} - \sum_{i=1}^{r} \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i = \mathbf{u} - \text{proj}_W \mathbf{u}$$

is orthogonal to each $\mathbf{v}_i$. i.e. $\mathbf{w}_2 \in W^{\perp}$.

(ii) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ are orthonormal basis for W, then the vector

$$\mathbf{w}_2 = \mathbf{u} - \sum_{i=1}^{r} \langle \mathbf{u}, \mathbf{v}_i \rangle \mathbf{v}_i = \mathbf{u} - \text{proj}_W \mathbf{u}$$

is orthonormal to each $\mathbf{v}_i$, that is, $\mathbf{w}_2 \in W^{\perp}$.

Theorem [Projection Theorem] Let W be a finite-dimensional subspace of an inner product space V. Then each $\mathbf{u}$ in V can be written uniquely in the form $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$ where $\mathbf{w}_1 = \text{proj}_W \mathbf{u} \in W$ and $\mathbf{w}_2 = \mathbf{u} - \text{proj}_W \mathbf{u} \in W^{\perp}$.

Theorem Every finite-dimensional inner product space has an orthonormal basis.

Gram–Schmidt Process Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an arbitrary basis for an inner product space V. Here the problem is to convert this arbitrary basis into an orthonormal basis for V.

The steps of Gram–Schmidt process to construct this orthonormal basis are given below.

Step 1 Take $\mathbf{w}_1 = \mathbf{v}_1$

Step 2 Compute $\mathbf{w}_2 = \mathbf{v}_2 - \dfrac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1$ which is orthogonal to $\mathbf{w}_1$.

Step 3 Compute $\mathbf{w}_3 = \mathbf{v}_3 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1 - \dfrac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2} \mathbf{w}_2$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2\}$.

Step 4 Compute $\mathbf{w}_4 = \mathbf{v}_4 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2} \mathbf{w}_2 - \dfrac{\langle \mathbf{v}_4, \mathbf{w}_3 \rangle}{\|\mathbf{w}_3\|^2} \mathbf{w}_3$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

… … … … … … … …

Step *n* Compute $\mathbf{w}_n = \mathbf{v}_n - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1 - \cdots - \dfrac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\|\mathbf{w}_{n-1}\|^2} \mathbf{w}_{n-1}$ which is orthogonal to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots, \mathbf{w}_{n-1}\}$.

Step (n + 1) The set $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots, \mathbf{w}_n\}$ is an orthogonal basis for V.

Step (n + 2) The set $\left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \cdots, \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|}\right\}$ is an orthonormal basis for V.

Theorem [The QR-Decomposition] If $\mathbf{A}$ is an $m \times n$ matrix with linearly independent columns, then the matrix $\mathbf{A}$ can be factored as $\mathbf{A} = \mathbf{QR}$ where $\mathbf{Q}$ is an $m \times n$ matrix whose columns form an orthonormal basis for columns space of $\mathbf{A}$, and $\mathbf{R}$ is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Steps to find the QR-decomposition of a matrix A
Let $\mathbf{A}$ be a matrix of order $m \times n$.

Step 1 Take $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ as column vectors of $\mathbf{A}$.

Step 2 Applying the Gram–Schmidt process on $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, find the orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ of the column space of $\mathbf{A}$.

Step 3 Make the matrix $\mathbf{Q}$ with columns $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$.

Step 4 Upper triangular matrix

$$\mathbf{R} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{q}_1 \rangle & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle & \cdots & \langle \mathbf{v}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2 \rangle & \cdots & \langle \mathbf{v}_n, \mathbf{q}_2 \rangle \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \vdots & \langle \mathbf{v}_n, \mathbf{q}_n \rangle \end{bmatrix}$$

Step 5 Matrix $\mathbf{A}$ factored as $\mathbf{A} = \mathbf{QR}$.

Theorem [The Best Approximation Theorem] If W is a finite-dimensional subspace of an inner product space V and $\mathbf{u}$ is any fixed vector outside the subspace W, then the orthogonal projection vector $\text{Proj}_W \mathbf{v}$ in the subspace W is a best approximation of $\mathbf{u}$ by the vector of subspace W.

$$\|\mathbf{v} - \text{proj}_W \mathbf{v}\| < \|\mathbf{v} - \mathbf{w}\| \qquad \text{for } \forall \mathbf{w} \in W \text{ and } \mathbf{w} \neq \text{proj}_W \mathbf{v}.$$

Error of Approximation If the vector $\text{proj}_W\mathbf{v}$ in subspace W is called *approximated error vector*, then $\|\mathbf{v} - \text{proj}_W \mathbf{v}\|$ is called an *error of approximation*.

Least-Squares Solution of the Inconsistent System If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b}$ is a vector in R^n, then a least-squares solution of the inconsistent system $\mathbf{Ax} = \mathbf{b}$ is $\hat{\mathbf{x}}$ in R^n such that $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\| \leq \|\mathbf{b} - \mathbf{Ax}\|$ for all $\mathbf{x} \in R^n$.

Theorem The least-squares solutions of the system of linear equation $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$ are the solutions of the normal equations $\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$.

Result If the columns of matrix $\mathbf{A}$ are linearly independent, then the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible.

Theorem [Uniqueness of Least-Squares Solution] If the columns of matrix $\mathbf{A}$ are linearly independent, then the system of linear equations $\mathbf{Ax} = \mathbf{b}$ has a unique least-squares solution $\hat{\mathbf{x}}$ and it is given by $\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$. (Note that vector $\mathbf{b} = \overline{\mathbf{b}}$).

5 Matrix Eigenvalue and Eigenvector Problems

In this chapter we will define two important concepts, that is, eigenvalues and eigenvectors of the square matrices. We will also see some of the uses of eigenvalues and eigenvectors in the diagonalization of the square matrix, quadratic form representation, etc.

5.1 EIGENVALUES AND EIGENVECTORS

Let us begin with the following definition.

Definition: Eigenvector of a Square Matrix

A vector $\mathbf{x} \in R^n$ is called an eigenvector of a square matrix $\mathbf{A} = [a_{jk}]$ of order n if

$$\mathbf{Ax} = \lambda \mathbf{x} \qquad \text{for some scalar } \lambda$$

where λ is called the *eigenvalue* of A.

Note $\mathbf{x}$ can be read as an *eigenvector* corresponding to λ. Geometrically, $\mathbf{x}$ is an eigenvector of R^n if the vector $\mathbf{Ax}$ lies on the same line as that of $\mathbf{x}$ passing through the origin.

EXAMPLE 5.1 Let $\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$; $\quad \mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Are $\mathbf{u}$ and $\mathbf{v}$ the eigenvectors of $\mathbf{A}$?

Solution: $\mathbf{Au} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix} \neq \lambda \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for any real λ.

Therefore $\mathbf{u}$ is not an eigenvector of $\mathbf{A}$.

$$\mathbf{Av} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

or $$\mathbf{Av} = 2\mathbf{v}$$

That is, $\mathbf{Av} = 2\mathbf{v}$ lies on the same line as that of **v** passing through the origin. So **v** is an eigenvectors of **A** and 2 is an eigenvalue of **A**. Moreover **A** only 'stretches' **v**. Therefore **v** is an eigenvector of **A**.

EXAMPLE 5.2 Let $\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$. Is **u** an eigenvector of **A**?

Solution: $$\mathbf{Au} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix} = -\begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

That is, $\mathbf{Au} = -\mathbf{u}$

Therefore **u** is an eigenvector of **A**. Moreover, **Au** is in the opposite direction to that of **u** and on the same line of **u** passing through the origin.

Characteristic Equation

Let **A** be a square matrix and λ be a scalar. Then the scalar equation,

$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$ where **I** is the identity matrix of same order as that of matrix **A**,

is called the *characteristic equation* of **A**.

Theorem 5.1 [Eigenvalues]

The eigenvalues of a square matrix **A** are the roots of the characteristic equation of **A**.

Remarks:

(i) A square matrix of order n has at least one eigenvalue and at most n numerically different eigenvalues.

(ii) If **x** is an eigenvector of a matrix **A** corresponding to an eigenvalue λ, then $k\mathbf{x}$ ($k \neq 0$) is also an eigenvector corresponding to the eigenvalue λ.

EXAMPLE 5.3 Find the characteristic equation of the matrix $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$. Also, calculate its eigenvalues and eigenvectors.

Solution: The characteristic equation of the given matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 3 & -1 \\ -2 & \lambda - 4 \end{bmatrix}\right) = 0$$

$$(\lambda - 3)(\lambda - 4) - 2 = 0$$

$$\lambda^2 - 7\lambda + 10 = 0 \qquad \text{(i)}$$

$$(\lambda - 5)(\lambda - 2) = 0$$

$$\therefore \qquad \lambda = 5 \text{ or } \lambda = 2$$

Equation (i) is called the characteristic equation of the matrix **A** and by Theorem 5.1, its roots 5 and 2 are the eigenvalues of the matrix **A**.

(i) *Eigenvector of* **A** *corresponding to the eigenvalue* $\lambda = 5$.
Suppose $x = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = 5$.

$$\therefore \qquad (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(5\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The solution is $\qquad 2x_1 - x_2 = 0 \quad \text{or} \quad x_2 = 2x_1$

$$\mathbf{x} = [x_1, x_2]$$

$$= [x_1, 2x_1]$$

$$= x_1[1, 2] \qquad (\text{if we choose } x_1 = 1)$$

Thus $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 5$.

(ii) *Eigenvector corresponding to the eigenvalue* $\lambda = 2$
Suppose $x = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = 2$.

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} -1 & -1 \\ -2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The solution is $\qquad x_1 + x_2 = 0 \quad \text{or} \quad x_2 = -x_1$

$$\therefore \qquad \mathbf{x} = [x_1, x_2]$$

$$= [x_1, -x_1]$$

$$= x_1[1, -1] \qquad (\text{if we choose } x_1 = 1)$$

Thus $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 2$.

EXAMPLE 5.4 Find the characteristic equation of the matrix $\mathbf{A} = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$. Also, calculate its eigenvalues and eigenvectors.

Solution: The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \mathbf{0}$$

$$\det\left(\lambda\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} - \begin{bmatrix}-1 & -2 & -2\\ 1 & 2 & 1\\ -1 & -1 & 0\end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix}\lambda+1 & 2 & 2\\ -1 & \lambda-2 & -1\\ 1 & 1 & \lambda\end{bmatrix}\right) = 0$$

$$(\lambda + 1)(\lambda^2 - 2\lambda + 1) - 2(1 - \lambda) + 2(1 - \lambda) = 0$$

$$\lambda^3 - \lambda^2 - \lambda + 1 = 0 \qquad \text{(i)}$$

$$(\lambda^2 - 1)(\lambda - 1) = 0$$

$$\therefore \qquad \lambda = -1, 1, 1$$

Thus, Eq. (i) is the characteristic equation of a matrix **A** and by Theorem 5.1, its roots 1, 1, –1 are the eigenvalues of matrix **A**.

(i) *Eigenvectors corresponding to the eigenvalue* $\lambda = 1$

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = 1$.

$$\therefore \qquad (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix}2 & 2 & 2\\ -1 & -1 & -1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

$$\sim \begin{bmatrix}1 & 1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \qquad \text{(after some elementary row operations)}$$

The solution is therefore given by

$$x_1 + x_2 + x_3 = 0$$

Let us choose $x_1 = s$, $x_2 = t$, $x_3 = -s - t$

$$\therefore \qquad \mathbf{x} = [x_1, x_2, x_3]$$

$$= [s, t, -s - t]$$

$$= s[1, 0, -1] + t[0, 1, -1] \qquad \text{(If we choose } s = 1, t = 0 \text{ and } s = 0 \text{ and } t = 1\text{)}$$

Therefore $\begin{bmatrix}1\\ 0\\ -1\end{bmatrix}$ and $\begin{bmatrix}0\\ 1\\ -1\end{bmatrix}$ are the eigenvectors of **A** corresponding to the eigenvalue $\lambda = 1$.

(ii) *Eigenvectors corresponding to the eigenvalue* $\lambda = -1$

Suppose $x = [x_1 \;\; x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = -1$.

$$(\lambda \mathrm{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(-\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 0 & 2 & 2 \\ -1 & -3 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{(after some elementary row operations)}$$

The solution is

$$x_2 + x_3 = 0; \qquad x_1 + 2x_2 = 0 \qquad \text{or} \qquad x_3 = -x_2; \qquad x_1 = -2x_2$$

$$\therefore \qquad \mathbf{x} = [x_1, x_2, x_3]$$

$$= [-2x_2, x_2, -x_2]$$

$$= x_2\,[-2, 1, -1] \qquad \text{(If we choose } x_2 = 1)$$

Therefore $\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$ is an eigenvector of matrix **A** corresponding to the eigenvalue $\lambda = -1$.

Theorem 5.2 [Eigenvalues and Eigenvectors]

Let **A** be a square matrix of order n and λ be a real number. Then the following statements are equivalent.

(i) λ is an eigenvalue of **A**

(ii) The system of equations $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ has non-trivial solutions.

(iii) There is a nonzero vector **x** in R^n suich that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

(iv) λ is a solution of the characteristic equation $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$.

Real Matrices with Complex Eigenvalues and Eigenvectors

Since real polynomials may have complex roots (which then occur in conjugate pairs), a real matrix may have complex eigenvalues and eigenvectors. See Example 5.5 below.

EXAMPLE 5.5 Find the eigenvalues and eigenvectors of the matrix $\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$, where $a \neq 0$.

Solution: Let the characteristic equation of the matrix **A** be

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-1 & 0 & 0 \\ 0 & \lambda-5 & 0 \\ 0 & 0 & \lambda-3 \end{bmatrix}\right) = 0$$

or $$(\lambda - 1)(\lambda - 5)(\lambda - 3) = 0$$

or $$\lambda = 1, 5, 3$$

Therefore the eigenvalues of the given diagonal matrix **C** are the diagonal entries 1, 5 and 3.

We have seen in the above examples that the diagonal entries are the eigenvalues. More generally, we have the following theorem.

Theorem 5.3 [Eigenvalues of a Lower Triangular, Upper Triangular or Diagonal Matrix]
If a square matrix **A** is a lower triangular (or upper triangular or diagonal) matrix, then the eigenvalues of **A** are the entries on the main diagonal of **A**.

Eigenspaces

If **A** is a square matrix of order n, then the solution space of the equation

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

is called the *eigenspace* of **A** corresponding to the eigenvalue λ.

Note: The eigenspace of a square matrix **A** of order n is a vector subspace of n dimensional space.

EXAMPLE 5.7 Find the eigenspace of the matrix $\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

Solution: The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-4 & 1 & -6 \\ -2 & \lambda-1 & -6 \\ -2 & 1 & \lambda-8 \end{bmatrix}\right) = 0$$

$$\lambda^3 - 13\lambda^2 + 40\lambda - 36 = 0$$

or $$(\lambda - 2)^2(\lambda - 9) = 0$$

$\therefore$ $$\lambda = 2, 2, 9$$

Therefore the roots 2, 2 and 9 are the eigenvalues of matrix **A**.

(i) *Eigenvectors corresponding to the eigenvalue* $\lambda = 2$

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = 2$.

$$\therefore \qquad (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(2\mathrm{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} -2 & 1 & -6 \\ -2 & 1 & -6 \\ -2 & 1 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} -2 & 1 & -6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{(after some elementary row operations)}$$

The solution is therefore given by

$$-2x_1 + x_2 - 6x_3 = 0$$

Let us choose $x_1 = s$, $x_3 = t$, $x_2 = 2s + 6t$

$$\mathbf{x} = [x_1, x_2, x_3]$$

$$= [s, 2s + 6t, t] \qquad \text{(If we choose } s = 1, t = 0 \text{ and } s = 0, t = 1\text{)}$$

$$= s\,(1, 2, 0) + t\,(0, 6, 1)$$

Therefore the eigenspace is a 2-dimensional subspace of R^3 generated by the basis $\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix}\right\}$

corresponding to the eigenvalue $\lambda = 2$. Hence the eigenspace $E(2) = \left\{ s\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \middle| \; s, t \in R \right\}$

(ii) *Eigenvectors corresponding to the eigenvalue* $\lambda = 9$

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = 9$.

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(9\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 5 & 1 & -6 \\ -2 & 8 & -6 \\ -2 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{(after some elementary row operations)}$$

The solution is therefore given by

$$-x_1 + x_2 = 0 \qquad -x_2 + x_3 = 0$$

$$\therefore \qquad \mathbf{x} = [x_1, x_2, x_3]$$

$$= [x_1, x_1, x_1]$$

$$= x_1\,[1, 1, 1]$$

Therefore the eigenspace is a 1-dimensional subspace of R^3 generated by the basis $\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ corresponding to the eigenvalue $\lambda = 9$. Hence the eigenspace $E(9) = \left\{ s\begin{bmatrix}1\\1\\1\end{bmatrix} \middle|\, s \in R \right\}$

EXAMPLE 5.8 Find the eigenspace of the matrix $\mathbf{A} = \begin{bmatrix}2 & 3\\3 & -6\end{bmatrix}$.

Solution: The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}2 & 3\\3 & -6\end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix}\lambda-2 & -3\\-3 & \lambda+6\end{bmatrix}\right) = 0$$

or
$$\lambda^2 + 4\lambda - 21 = 0 \tag{i}$$

$$(\lambda + 7)(\lambda - 3) = 0$$

$$\therefore \qquad \lambda = -7 \qquad \text{or} \qquad \lambda = 3$$

Therefore the roots –7, 3 are the complex eigenvalues of the matrix **A**.

(i) *Eigenvectors corresponding to the eigenvalue* $\lambda = -7$.
Suppose $\mathbf{x} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = -7$.

$$\therefore \qquad (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(-7\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix}-9 & -3\\-3 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$\sim \quad \begin{bmatrix}0 & 0\\-3 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \qquad \text{(after some elementary row operations)}$$

The solution is therefore given by

$$-3x_1 - x_2 = 0$$

$$\mathbf{x} = [x_1, x_2]$$

$$= [x_1, -3x_1]$$

$$= x_1[1, -3]$$

Thus $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ is an eigenvector of a matrix **A** corresponding to the eigenvalue $\lambda = -7$. Therefore the eigenspace is a 1-dimensional subspace of R^2 generated by the basis $\left\{\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right\}$ corresponding to the eigenvalue $\lambda = -7$. Hence eigenspace

$$E(-7) = \left\{ s\begin{bmatrix} 1 \\ -3 \end{bmatrix} \middle| \, s \in R \right\}$$

(ii) *Eigenvectors corresponding to the eigenvalue* $\lambda = 3$.
Suppose $x = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = 3$.

$$\therefore \quad (\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(3\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{(after some elementary row operations)}$$

The solution is therefore given by

$$x_1 - 3x_2 = 0$$

$$\therefore \quad \mathbf{x} = [x_1, x_2]$$

$$= [3x_2, x_2]$$

$$= x_2[3, 1]$$

Thus $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector of the matrix **A** corresponding to the eigenvalue $\lambda = 3$. Therefore the eigenspace is a 1-dimensional subspace of R^2 generated by the basis $\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}$ corresponding to the eigenvalue $\lambda = 3$. Hence eigenspace

$$E(3) = \left\{ s\begin{bmatrix} 3 \\ 1 \end{bmatrix} \middle| \, s \in R \right\}$$

The following theorem describes some good properties of the eigenvalues of the square matrix.

Theorem 5.4

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ (may not be distinct) are the eigenvalues of a square matrix of order n, then:

(i) The determinant of **A**, i.e. det $\mathbf{A} = \lambda_1 \lambda_2 \ldots \lambda_n$

(ii) The trace of **A**, i.e. tr $\mathbf{A} = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Corollary 5.1 A square matrix **A** is invertible if and only if $\lambda = 0$ is not an eigenvalue of **A**.

Corollary 5.2 If the characteristic equation of **A** is $\lambda^n + C_1 \lambda^{n-1} + \cdots + C_n$, then det $\mathbf{A} = (-1)^n C_n$.

Remark From the above two corollaries, a square matrix **A** is invertible if and only if $C_n \neq 0$.

EXAMPLE 5.9 Find the eigenvalues of the matrix $\mathbf{A} = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{bmatrix}$ and verify the results of Theorem 5.4.

Solution: The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda & 0 & -3 \\ 0 & \lambda - 3 & 0 \\ -3 & 0 & \lambda \end{bmatrix}\right) = 0$$

or $$(\lambda - 3)(\lambda^2 - 9) = 0$$

$\therefore$ $$\lambda = 3, 3, -3$$

Therefore the roots $\lambda_1 = 3$, $\lambda_2 = 3$, and $\lambda_3 = -3$ are the eigenvalues of matrix **A**. Now we will verify the results of Theorem 5.4.

(i) $\det \mathbf{A} = \begin{vmatrix} 0 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{vmatrix} = -27$

Now $\lambda_1 \lambda_2 \lambda_3 = (3)(3)(-3) = -27$

Hence det $\mathbf{A} = \lambda_1 \lambda_2 \lambda_3$.

(ii) tr $\mathbf{A} = a_{11} + a_{22} + a_{33} = 0 + 3 + 0 = 3$

Now $\lambda_1 + \lambda_2 + \lambda_3 = 3 + 3 + (-3) = 3$

Hence tr $\mathbf{A} = \lambda_1 + \lambda_2 + \lambda_3$.

EXAMPLE 5.10 Find the eigenvalues of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and verify Corollary 5.1.

Solution: The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 1 & -3 & -4 \\ 0 & \lambda - 1 & -1 \\ 0 & 0 & \lambda \end{bmatrix}\right) = 0$$

or $$(\lambda - 1)(\lambda - 1)(\lambda) = 0$$

∴ $$\lambda = 1, 1, 0$$

Therefore the roots $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 1$ are the eigenvalues of matrix **A**. By Theorem 5.4,

$$\det \mathbf{A} = \lambda_1 \lambda_2 \lambda_3 = (0)(1)(1) = 0.$$

Thus the given matrix **A** is not invertible.

Theorem 5.5

If λ is an eigenvalue of a square matrix **A** corresponding to an eigenvector **x**, then for a positive integer k, λ^k is an eigenvalue of $\mathbf{A}^k$ corresponding to an eigenvector **x**.

Proof: λ is an eigenvalue of **A** corresponding to an eigenvector **x**. Hence

$$\mathbf{Ax} = \lambda \mathbf{x}$$

Now
$$\begin{aligned} \mathbf{A}^k\mathbf{x} &= \mathbf{A}^{k-1}(\mathbf{Ax}) \\ &= \mathbf{A}^{k-1}(\lambda \mathbf{x}) \\ &= \lambda \mathbf{A}^{k-2}(\mathbf{Ax}) \\ &= \lambda A^{k-2}(\lambda \mathbf{x}) \\ &= \lambda^2 \mathbf{A}^{k-2}(\mathbf{x}) \\ &\dots\dots \quad \dots \\ &= \lambda^{k-1}(\mathbf{Ax}) \\ &= \lambda^{k-1}(\lambda \mathbf{x}) \\ &= \lambda^k \mathbf{x} \end{aligned}$$

Therefore λ^k is an eigenvalue of $\mathbf{A}^k$ corresponding to an eigenvector **x**.

EXAMPLE 5.11 Find the eigenvalues and eigenvectors of $\mathbf{A}^4$ for $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

EXAMPLE 5.13 If $\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$, find the eigenvalues and the bases for the eigenspaces of

(i) $\mathbf{A}^{-1}$ (ii) $\mathbf{A} - 5\mathbf{I}$ (iii) $\mathbf{A} + 4\mathbf{I}$

Solution: We showed in Example 5.7 that $\mathbf{B}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right\}$ and $\mathbf{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ are the bases for eigenspace of **A** corresponding to eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 9$ respectively.

(i) By using the Result 2, $\mathbf{B}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right\}$ and $\mathbf{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ are the bases for eigenspace of $\mathbf{A}^{-}$ corresponding to eigenvalues $\dfrac{1}{\lambda_1} = \dfrac{1}{2}$ and $\dfrac{1}{\lambda_2} = \dfrac{1}{9}$ respectively.

(ii) By using the Result 2, $\mathbf{B}_1 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right\}$ is the basis for eigenspace of **A** corresponding to eigenvalue $\lambda_1 = 2$, hence $\mathbf{B}_1$ is the basis for eigenspace of $\mathbf{A} - 5\mathbf{I}$ corresponding to eigenvalue $\lambda_1 - 5 = 2 - 5 = -3$. Similarly, since $\mathbf{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is the basis for eigenspace of **A** corresponding to eigenvalues $\lambda_1 = 9$, hence $\mathbf{B}_2$ is the basis for eigenspace of $\mathbf{A} - 5\mathbf{I}$ corresponding to eigenvalue $\lambda_2 - 5 = 9 - 5 = 4$.

(iii) From the Result 2, since $\mathbf{B}_1$ and $\mathbf{B}_2$ are the bases for eigenspace of **A** corresponding to eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 9$ respectively, hence $\mathbf{B}_1$ and $\mathbf{B}_2$ are the bases for eigenspace of $\mathbf{A} + 4\mathbf{I}$ corresponding to eigenvalues $\lambda_1 + 4 = 2 + 4 = 6$ and $\lambda_2 + 4 = 9 + 4 = 13$ respectively.

Theorem 5.6 [Cayley–Hamilton Theorem]

Every square matrix **A** satisfies its characteristic equation. That is, if

$$a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = 0$$

is the characteristic equation of **A**, then

$$a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + \ldots + a_n\mathbf{A}^n = \mathbf{0}.$$

EXAMPLE 5.14 Verify the Cayley–Hamilton Theorem for $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 5 & 4 \end{bmatrix}$.

Solution: The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\begin{bmatrix}\lambda-1 & -3\\ -5 & \lambda-4\end{bmatrix}=0$$

$$(\lambda-1)(\lambda-4)-15=0$$

or
$$\lambda^2-5\lambda-11=0 \tag{i}$$

Equation (i) is the characteristic equation of **A**. To verify the Cayley–Hamilton Theorem, we have to show that $\mathbf{A}^2-5\mathbf{A}-11\mathbf{I}=\mathbf{0}$.

$$\mathbf{A}^2=\begin{bmatrix}1 & 3\\ 5 & 4\end{bmatrix}\begin{bmatrix}1 & 3\\ 5 & 4\end{bmatrix}=\begin{bmatrix}16 & 15\\ 25 & 31\end{bmatrix}$$

$$5\mathbf{A}=5\begin{bmatrix}1 & 3\\ 5 & 4\end{bmatrix}=\begin{bmatrix}5 & 15\\ 25 & 20\end{bmatrix}$$

$$11\mathbf{I}=11\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}11 & 0\\ 0 & 11\end{bmatrix}$$

$$\mathbf{A}^2-5\mathbf{A}-11\mathbf{I}=\begin{bmatrix}16 & 15\\ 25 & 31\end{bmatrix}-\begin{bmatrix}5 & 15\\ 25 & 20\end{bmatrix}-\begin{bmatrix}11 & 0\\ 0 & 11\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=\mathbf{0}$$

Hence **A** satisfies its characteristic equation.

EXAMPLE 5.15 Find $\mathbf{A}^2$, $\mathbf{A}^{-1}$ for $\mathbf{A}=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}$.

Solution: The characteristic equation of **A** is

$$\det(\lambda \mathrm{I}-\mathbf{A})=0$$

$$\det\begin{bmatrix}\lambda-1 & 0\\ -1 & \lambda-1\end{bmatrix}=0$$

$$\lambda^2-2\lambda+1=0 \tag{i}$$

Equation (i) is the characteristic equation of **A**. By the Cayley–Hamilton theorem, we have

$$\mathbf{A}^2-2\mathbf{A}+\mathbf{I}=\mathbf{0}$$

(i) $\mathbf{A}^2=2\mathbf{A}-\mathbf{I}$

$$=2\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}-\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}$$

(ii) Since $\mathbf{A}^2-2\mathbf{A}+\mathbf{I}=\mathbf{0}$

$$\mathbf{A}-2\mathbf{I}+\mathbf{A}^{-1}=\mathbf{0}$$

$$\mathbf{A}^{-1} = 2\mathbf{I} - \mathbf{A}$$

$$= 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

EXAMPLE 5.16 Express $\mathbf{A}^3$ as a linear combination of $\mathbf{I}$, $\mathbf{A}$, $\mathbf{A}^2$ for $\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Does $\mathbf{A}^{-1}$ exist?

Solution: The characteristic equation of $\mathbf{A}$ is

$$\det(\lambda \mathbf{I} - A) = 0$$

$$\det\begin{bmatrix} \lambda & 1 & 0 \\ -1 & -1 & 0 \\ 0 & -1 & \lambda \end{bmatrix} = 0$$

or
$$\lambda^3 - \lambda^2 + \lambda = 0 \tag{i}$$

Equation (i) is the characteristic equation of $\mathbf{A}$. By the Cayley–Hamilton theorem, we have

$$\mathbf{A}^3 - \mathbf{A}^2 + \mathbf{A} = \mathbf{0}$$

or
$$\mathbf{A}^3 = \mathbf{A}^2 - \mathbf{A}$$

Since $\lambda = 0$ is an eigenvalue of $\mathbf{A}$, $\mathbf{A}^{-1}$ does not exist.

EXERCISE SET 1

1. Find the characteristic equation of the following matrices:

(i) $\begin{bmatrix} \pi & 1 \\ 0 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix}$ (iii) $\begin{bmatrix} 2 & -2 & 2 \\ 0 & 1 & 1 \\ -4 & 8 & 3 \end{bmatrix}$

(iv) $\begin{bmatrix} 3 & -3 & 2 \\ 1 & -1 & 2 \\ 1 & -3 & 4 \end{bmatrix}$ (v) $\begin{bmatrix} 17 & -10 & -5 \\ 45 & -28 & -15 \\ -30 & 20 & 12 \end{bmatrix}$ (vi) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

2. Find the eigenvalues and the bases for eigenspaces of the matrices in the above Example 1.

3. Find the eigenvalues of the following matrices by the method of inspection:

(i) $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

4. Find the bases for eigenspaces of the matrices in the above Example 3.

5. Find the det **A** whose characteristic polynomial $p(\lambda)$ is as follows:

(i) $p(\lambda) = \lambda^{10} - \lambda^7 + 5\lambda^6 + \lambda^4 - 2\lambda^3 - 14$

(ii) $p(\lambda) = \lambda^7 + 4\lambda^5 - \lambda^4 + 3\lambda^3 - \lambda^2 + \lambda + 7$

(iii) $p(\lambda) = \lambda^2 + \lambda$

6. Determine a, b, c, d, e, f given that the vectors [1, 1, 1], [1, 0, –1] and [1, –1, 0] are the given vectors of the matrix $\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}$.

7. Find the eigenvalues, eigenvectors of $\mathbf{A}^5$ for $\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$.

8. If $\mathbf{A} = \begin{bmatrix} -2 & 2 & 3 \\ -2 & 3 & 2 \\ -4 & 2 & 5 \end{bmatrix}$, find the eigenvalues and the bases for the eigenspaces of:

(i) $\mathbf{A}^{-1}$ (ii) $\mathbf{A} - 5\mathbf{I}$ (iii) $\mathbf{A} + 4\mathbf{I}$

9. Verify the Cayley–Hamilton theorem for the following matrices:

(i) $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (ii) $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

10. Express $\mathbf{A}^{-1}$, $\mathbf{A}^4$ and $\mathbf{A}^3$ as a linear combination of $\mathbf{A}^2$, $\mathbf{A}$, $\mathbf{I}$ for $\mathbf{A} = \begin{bmatrix} 5 & 4 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$.

11. By using the Cayley–Hamilton theorem, find $\mathbf{A}^{-1}$ for the matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$.

12. Find the eigenvalues and eigenvectors for $\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

13. Suppose that **A** has eigenvalues 1, 2, 4. What is the trace of $\mathbf{A}^2$? What is the determinant of $\mathbf{A}^{-1}$?

5.2 EIGENVALUES OF REAL AND COMPLEX MATRICES

In this section we will define some special types of real and complex matrices. Also, we will discuss their eigenvalues.

Real Matrices

Let us start with basic definitions

Definition: Symmetric Matrix

If $\mathbf{A} = [a_{ij}]$ is a real $n \times n$ square matrix and

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A} \qquad \text{that is} \qquad [a_{ij}] = [a_{ji}], \text{ for each pair } (i, j)$$

then **A** is called a *symmetric matrix*.

For example, If $\mathbf{A} = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 4 & 6 \\ -3 & 6 & 8 \end{bmatrix}$ then $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} -1 & 2 & -3 \\ 2 & 4 & 6 \\ -3 & 6 & 8 \end{bmatrix} = \mathbf{A}$. Therefore the given matrix is a symmetric matrix.

Definition: Skew-symmetric Matrix

If $\mathbf{A} = [a_{ij}]$ is a real $n \times n$ square matrix and

$$\mathbf{A}^{\mathrm{T}} = -\mathbf{A} \qquad \text{that is} \qquad [a_{ij}] = -[a_{ji}], \text{ for each pair } (i, j)$$

then **A** is called a *skew-symmetric matrix*.

For example If $\mathbf{A} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$ then $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix} = -\mathbf{A}$

Therefore the given matrix is a *skew-symmetric matrix*.

Remarks:

(i) The diagonal entries of the skew-symmetric matrix are always zero.

(ii) Every real square matrix **A** can be expressed as the sum of a symmetric matrix and a skew-symmetric matrix, that is, every real square matrix **A** can be expressed as

$$\mathbf{A} = \mathbf{B} + \mathbf{C}$$

where $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$ is symmetric matrix

and $\mathbf{C} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathrm{T}})$ is a skew-symmetric matrix.

EXAMPLE 5.17 Express the square matrix $\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \\ 8 & 3 & 6 \end{bmatrix}$ as the sum of a symmetric matrix and a skew-symmetric matrix.

Solution: Here

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \\ 8 & 3 & 6 \end{bmatrix}$$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 8 \\ 3 & -1 & 3 \\ 5 & 4 & 6 \end{bmatrix}$$

Take

$$\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}}) = \frac{1}{2}\begin{bmatrix} 2 & 5 & 13 \\ 5 & -2 & 7 \\ 13 & 7 & 12 \end{bmatrix} = \begin{bmatrix} 1 & \frac{5}{2} & \frac{13}{2} \\ \frac{5}{2} & -1 & \frac{7}{2} \\ \frac{13}{2} & \frac{7}{2} & 6 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^{\mathrm{T}}) = \frac{1}{2}\begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & 1 \\ 3 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 0 \end{bmatrix}$$

One can easily show that **B** is a symmetric matrix and **C** is a skew-symmetric matrix. The next theorem will describe the character of eigenvalues of symmetric and skew-symmetric matrices.

Theorem 5.7 [Eigenvalues of Real Symmetric Matrices]

The eigenvalues of a real symmetric matrix are real and the eigenvectors from different eigenspaces are orthogonal.

Theorem 5.8 [Eigenvalues of Real Skew-Symmetric Matrices]

The eigenvalues of a real skew-symmetric matrix are pure imaginary or zero.

EXAMPLE 5.18 Let $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Show that **A** is a symmetric matrix and also verify Theorem 5.7.

Solution: Here $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

and

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = \mathbf{A}$$

Therefore **A** is a symmetric matrix. To verify Theorem 5.7, we have to find the eigenvalues of **A**. The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-3 & -1 \\ -1 & \lambda-3 \end{bmatrix}\right) = 0$$

$$\lambda^2 - 6\lambda + 8 = 0 \qquad \text{(i)}$$

or
$$(\lambda - 4)(\lambda - 2) = 0$$

∴
$$\lambda = 4 \quad \text{or} \quad \lambda = 2$$

Therefore, a real symmetric matrix **A** has real eigenvalues $\lambda_1 = 4$, $\lambda_2 = 2$.

***Case* (i)** *Eigenvectors corresponding to the eigenvalue* $\lambda_1 = 4$

Suppose $\mathbf{x} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 4$.
Therefore,

$$(\lambda_1\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(4\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 - x_2 = 0; \qquad x_2 = x_1$$

$$\begin{aligned} \mathbf{x} &= [x_1, x_2] \\ &= [x_1, x_1 \\ &= x_1[1, 1] \end{aligned}$$

Thus, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 4$.

***Case* (ii)** *Eigenvectors corresponding to the eigenvalue* $\lambda_2 = 2$

Suppose $\mathbf{x} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.
Therefore,

$$(\lambda_2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 + x_2 = 0; \qquad x_2 = -x_1$$

$$\begin{aligned} \mathbf{x} &= [x_1, x_2] \\ &= [x_1, -x_1] \\ &= x_1[1, -1] \end{aligned}$$

Thus, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 2$.

With respect to the Euclidean inner product, $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. Therefore, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal eigenvectors.

EXAMPLE 5.19 Let $\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix}$. Show that **A** is a symmetric matrix and also verify Theorem 5.7.

Solution: Here $\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix}$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix} = \mathbf{A}$$

Therefore, **A** is a symmetric matrix. To verify Theorem 5.7., we need the eigenvalues of **A**. The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 2 & -1 & -3 \\ -1 & \lambda - 2 & -3 \\ -3 & -3 & \lambda - 20 \end{bmatrix}\right) = 0$$

$$(\lambda - 2)(\lambda^2 - 22\lambda + 31) + 1(-\lambda + 20 - 9) - 3(3 + 3\lambda - 6) = 0$$

$$\lambda^3 - 24\lambda^2 + 65\lambda - 42 = 0$$

or $$(\lambda - 1)(\lambda - 2)(\lambda - 21) = 0$$

$\therefore$ $$\lambda = 1, 2, 21$$

Therefore, a real symmetric matrix **A** has real eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 21$.

***Case* (i)** *Eigenvectors corresponding to the eigenvalue* $\lambda_1 = 1$.

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$. Therefore,

$$(\lambda_1 \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} -1 & -1 & -3 \\ -1 & -1 & -3 \\ -3 & -3 & -19 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1 + x_2 = 0; \qquad x_3 = 0$$

$$\begin{aligned}\mathbf{x} &= [x_1, -x_1, x_3]\\ &= [x_1, -x_1, 0]\\ &= x_1[1, -1, 0]\end{aligned}$$

Thus, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$.

***Case* (ii)** *Eigenvectors corresponding to the eigenvalue* $\lambda_2 = 2$.

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$. Therefore,

$$(\lambda_2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 0 & -1 & -3 \\ -1 & 0 & -3 \\ -3 & -3 & -18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 & -3 \\ -1 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_2 + 3x_3 = 0; \qquad x_1 + 3x_3 = 0$$

$$\begin{aligned}\mathbf{x} &= [x_1, x_2, x_3]\\ &= [-3x_3, -3x_3, x_3]\\ &= x_3[-3, -3, 1]\end{aligned}$$

Thus, $v_2 = \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.

***Case* (iii)** *Eigenvectors corresponding to the eigenvalue* $\lambda_3 = 21$.

Suppose $\mathbf{x} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to eigenvalue $\lambda_3 = 21$. Therefore,

$$(\lambda_3\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$(21\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} 19 & -1 & -3 \\ -1 & 19 & -3 \\ -3 & -3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -19 & 3 \\ -1 & 19 & -3 \\ -3 & -3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -19 & 3 \\ 0 & 0 & 0 \\ 0 & -60 & 10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -19 & 3 \\ 0 & 0 & 0 \\ 0 & -6 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -6 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1 - x_2 = 0; \qquad x_3 - 6x_2 = 0$$

$$\begin{aligned} \mathbf{x} &= [x_1, x_2, x_3] \\ &= [x_2, x_2, 6x_2] \\ &= x_2[1, 1, 6] \end{aligned}$$

Thus, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_3 = 22$.

With respect to the Euclidean inner product, $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for each i and j where $i, j = 1, 2, 3$. Therefore $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are orthogonal eigenvectors.

EXAMPLE 5.20 Let $\mathbf{A} = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$. Show that **A** is skew symmetric and also verify Theorem 5.8.

Solution: Here $\mathbf{A} = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} = -\mathbf{A}$$

Therefore **A** is a skew-symmetric matrix. To verify Theorem 5.8, we want the eigenvalues of **A**. The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda & b \\ -b & \lambda \end{bmatrix}\right) = 0$$

or $$\lambda^2 + b^2 = 0$$

∴ $$\lambda = \pm bi$$

Therefore a real skew-symmetric matrix **A** has pure imaginary eigenvalues $\lambda_1 = bi$, $\lambda_2 = -bi$.

EXAMPLE 5.21 Show that $\mathbf{A} = \begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}$ is skew-symmetric and also verify Theorem 5.8.

Solution: Here $\mathbf{A} = \begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 0 & -9 & 12 \\ 9 & 0 & -20 \\ -12 & 20 & 0 \end{bmatrix} = -\mathbf{A}$$

Therefore **A** is a skew-symmetric matrix. The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\begin{bmatrix} \lambda & -9 & 12 \\ 9 & \lambda & -20 \\ -12 & 20 & \lambda \end{bmatrix}\right) = 0$$

$$\lambda^3 + 625\lambda = 0$$

or $$\lambda(\lambda^2 + 625) = 0$$

∴ $$\lambda = 0 \quad \text{or} \quad \lambda = \pm 5i$$

Therefore, the given real skew-symmetric matrix **A** has pure imaginary eigenvalues $\lambda_1 = 5i$, $\lambda_2 = -5i$ and $\lambda_3 = 0$. Hence Theorem 5.8 is verified.

Orthogonal Matrix

If **A** is a real square matrix and

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I} \qquad \text{or} \qquad \mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1},$$

then **A** is called an *orthogonal matrix*.

Result The value of determinant of an orthogonal matrix is 1 and –1.

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$$

$$\det \mathbf{A}\mathbf{A}^{\mathrm{T}} = \det \mathbf{I}$$

$$\det \mathbf{A} \det \mathbf{A}^{\mathrm{T}} = 1$$

$$\det \mathbf{A} \det \mathbf{A} = 1$$

$$(\det \mathbf{A})^2 = 1$$

$$\therefore \qquad \det \mathbf{A} = \pm 1$$

For example, take
$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

then
$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Since det $\mathbf{A} = 1 \neq 0$, so $\mathbf{A}$ is an invertible matrix.

Hence
$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \operatorname{adj} \mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Therefore $\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}$. Hence $\mathbf{A}$ is an orthogonal matrix.

Theorem 5.9 [Orthogonal Matrix]

A real square matrix is orthogonal if and only if the rows of the matrix form an orthonormal set of vectors, or if and only if the columns of the matrix form an orthonormal set of vectors.

The following theorem will give the character of eigenvalues of orthogonal matrix.

Theorem 5.10 [Eigenvalues of Orthogonal Matrices]

The eigenvalues of an orthogonal matrix are real or complex conjugate in pairs and have the absolute value 1.

EXAMPLE 5.22 Show that the matrix $\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$ is an orthogonal matrix. Also verify the results of Theorem 5.9 and Theorem 5.10.

Solution: Here $\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{AA}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\therefore$ $\quad\mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$

Therefore **A** is an orthogonal matrix. Moreover the set $\left\{\left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right], [0, 1, 0], \left[\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right]\right\}$ of rows vectors of **A** is an orthogonal set. To verify the result of Theorem 5.10, we need the eigenvalues of **A**.

The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\begin{bmatrix} \lambda - \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \lambda - 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \lambda + \frac{1}{\sqrt{2}} \end{bmatrix}\right) = 0$$

or $\quad(\lambda - 1)(\lambda^2 - 1) = 0$

$\therefore$ $\quad\lambda = 0 \quad \text{or} \quad \lambda = 1 \quad \text{or} \quad \lambda = -1$

Therefore the matrix **A** has real eigenvalues $\lambda = 0$ or $\lambda = 1$ or $\lambda = -1$ with the absolute value 1.

EXAMPLE 5.23 Show that the matrix $\mathbf{A} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$ is an orthogonal matrix. Also, verify the results of Theorem 5.9 and Theorem 5.10.

Solution: Here $\mathbf{A} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$$

$$\mathbf{AA}^{\mathrm{T}} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}\begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\therefore \qquad \mathbf{A}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$$

Therefore **A** is an orthogonal matrix. The set $\left\{\left[\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right]\right\}$ is an orthogonal set of vectors. To verify the result of Theorem 5.10, we have to find the eigenvalues of **A**. The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\begin{bmatrix} \lambda - \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \lambda - \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} & \lambda + \frac{2}{3} \end{bmatrix}\right) = 0$$

$$(\lambda + 1)\left(\lambda^2 - \frac{5}{3}\lambda + 1\right) = 0$$

$$\therefore \qquad \lambda = -1 \quad \text{or} \quad \lambda = \frac{5 \pm i\sqrt{11}}{6}$$

Therefore, the orthogonal matrix **A** has real eigenvalue $\lambda_1 = -1$ and complex conjugates pair $\lambda_2 = \dfrac{5+\sqrt{11}i}{6}$ and $\lambda_3 = \dfrac{5-i\sqrt{11}}{6}$. We can also show that $|\lambda_1| = |\lambda_2| = |\lambda_3| = 1$.

Complex Matrices

A matrix with complex entries is called *complex matrix*. Here we will discuss some special complex matrices: Hermitian, skew-Hermitian, unitary and normal matrices.

Conjugate of Complex Matrix

If $\mathbf{A} = [a_{ij}]$ is a complex matrix, then the conjugate of **A** is $\overline{\mathbf{A}} = [\overline{a}_{ij}]$ where $[\overline{a}_{ij}]$ is a complex conjugate of **A**. For example, if

$$\mathbf{A} = \begin{bmatrix} i & 2+i \\ 4 & 3-2i \end{bmatrix}$$

then

$$\overline{\mathbf{A}} = \begin{bmatrix} -i & 2-i \\ 4 & 3+2i \end{bmatrix}$$

Hermitian Matrix

If $\mathbf{A} = [a_{ij}]$ is a complex matrix and

$$\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A} \qquad \text{where } \overline{\mathbf{A}}^{\mathrm{T}} = \text{conjugate transpose of } \mathbf{A},$$

that is, $a_{ij} = \overline{a}_{ji}$ for all i and j,
then **A** is called the Hermitian matrix. For example, if

(i) $$\mathbf{A} = \begin{bmatrix} 6 & 3+i \\ 3-i & 4 \end{bmatrix}$$

then $$\overline{\mathbf{A}} = \begin{bmatrix} 6 & 3-i \\ 3+i & 4 \end{bmatrix}$$

and $$\overline{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} 6 & 3+i \\ 3-i & 4 \end{bmatrix}$$

Hence $\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A}$

Thus **A** is a Hermitian matrix.

(ii) For $$\mathbf{B} = \begin{bmatrix} 4 & 1+2i & -i \\ 1-2i & -3 & 4+5i \\ i & 4-5i & 0 \end{bmatrix}$$

$$\overline{\mathbf{B}} = \begin{bmatrix} 4 & 1-2i & i \\ 1+2i & -3 & 4-5i \\ -i & 4+5i & 0 \end{bmatrix}$$

$$\overline{\mathbf{B}}^{\mathrm{T}} = \begin{bmatrix} 4 & 1+2i & -i \\ 1-2i & -3 & 4+5i \\ i & 4-5i & 0 \end{bmatrix}$$

$$\therefore \qquad \overline{\mathbf{B}}^{\mathrm{T}} = \mathbf{B}$$

Thus **B** is a Hermitian matrix.

Skew-Hermitian Matrix

If $\mathbf{A} = [a_{ij}]$ is a complex matrix and

$$\overline{\mathbf{A}}^{\mathrm{T}} = -\mathbf{A} \qquad \text{where } \overline{\mathbf{A}}^{\mathrm{T}} = \text{conjugate transpose of } \mathbf{A},$$

that is, $\quad a_{ij} = -\overline{a}_{ji} \quad$ for all i and j,

then **A** is called the skew- Hermitian matrix. For example, if

$$\mathbf{A} = \begin{bmatrix} i & 2+3i \\ -2+3i & -3i \end{bmatrix}$$

then

$$\overline{\mathbf{A}} = \begin{bmatrix} -i & 2-3i \\ -2-3i & 3i \end{bmatrix}$$

$$\overline{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} -i & -2-3i \\ 2-3i & 3i \end{bmatrix} = -\begin{bmatrix} i & 2+3i \\ -2+3i & -3i \end{bmatrix}$$

$$\therefore \qquad \overline{\mathbf{A}}^{\mathrm{T}} = -\mathbf{A}$$

Thus **A** is a skew-Hermitian matrix.

Remarks:

(i) If **A** is a Hermitian matrix, then the entries on the main diagonal must satisfy $a_{ii} = \overline{a}_{ii}$, that is, they are real.

(ii) If **A** is a skew-Hermitian matrix, then the entries on the main diagonal must satisfy $a_{ii} = -\overline{a}_{ii}$, that is, they are purely imaginary.

(iii) If **A** is a square matrix, then **A** can be written as **B** + **C** where $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \overline{\mathbf{A}}^{\mathrm{T}})$ is a real symmetric matrix and $\mathbf{C} = \frac{1}{2}(\mathbf{A} - \overline{\mathbf{A}}^{\mathrm{T}})$ is a real skew-symmetric matrix.

The following theorems will discuss the eigenvalues of Hermitian and skew-Hermitian matrices.

Theorem 5.11 [Eigenvalues of Hermitian Matrices]
The eigenvalues of a Hermitian matrix are real.

Theorem 5.12 [Eigenvalues of Skew-Hermitian Matrices]
The eigenvalues of a skew-Hermitian matrix are pure imaginary or zero.

$$\det\left(\begin{bmatrix} \lambda-1 & -i & -2 \\ i & \lambda & -1+i \\ -2 & -1-i & \lambda-3 \end{bmatrix}\right) = 0$$

or $$(\lambda+1)(\lambda^2 - 5\lambda + 1) = 0$$

∴ $$\lambda = -1 \quad \text{or} \quad \lambda = \frac{5 \pm \sqrt{21}}{2}$$

Therefore the eigenvalues are real.

EXAMPLE 5.26 Show that $\mathbf{A} = \begin{bmatrix} i & 2+i \\ -2+i & 0 \end{bmatrix}$ is a skew-Hermitian matrix and verify Theorem 5.12.

Solution: Here $\mathbf{A} = \begin{bmatrix} i & 2+i \\ -2+i & 0 \end{bmatrix}$

$$\overline{\mathbf{A}} = \begin{bmatrix} -i & 2-i \\ -2-i & 0 \end{bmatrix}$$

$$\overline{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} -i & -2-i \\ 2-i & 0 \end{bmatrix} = -\begin{bmatrix} i & 2+i \\ -2+i & 0 \end{bmatrix}$$

∴ $$\overline{\mathbf{A}}^{\mathrm{T}} = -\mathbf{A}$$

Thus **A** is a skew-Hermitian matrix. To verify Theorem 5.12, we will find the eigenvalues of **A**. The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\begin{bmatrix} \lambda-i & -2-i \\ 2-i & \lambda \end{bmatrix}\right) = 0$$

or $$\lambda^2 - \lambda i + 5 = 0$$

∴ $$\lambda = (1 \pm \sqrt{21})i$$

Thus the skew-Hermitian matrix **A** has purely imaginary eigenvalues $\lambda_1 = (1+\sqrt{21})\,i$ or $\lambda_2 = (1-\sqrt{21})\,i$

Unitary Matrix

If $\mathbf{A} = [a_{ij}]$ is a complex matrix and

$$\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A}^{-1} \qquad \text{or} \qquad \mathbf{A}\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{I}$$

where $\overline{\mathbf{A}}^{\mathrm{T}}$ = conjugate transpose of **A**, then A is called the *unitary matrix*. For example, if

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \sqrt{\frac{3}{4}}\,i \\ \sqrt{\frac{3}{4}}\,i & \frac{1}{2} \end{bmatrix}$$

EXAMPLE 5.28 Show that the matrix $\mathbf{A} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix}$ is a unitary matrix and the absolute values of eigenvalues of **A** are 1.

Solution: Here $\mathbf{A} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix}$

$$\bar{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix}$$

$$\mathbf{A}\bar{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix}\begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{A}\bar{\mathbf{A}}^{\mathrm{T}} = \mathbf{I}$$

Thus **A** is a unitary matrix. The characteristic equation **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\begin{bmatrix} \lambda - i & 0 & 0 \\ 0 & \lambda & -i \\ 0 & -i & \lambda \end{bmatrix} = 0$$

or $$(\lambda - i)(\lambda^2 + 1) = 0$$

$\therefore$ $$\lambda = i \quad \text{or} \quad \lambda = i \quad \text{or} \quad \lambda = -i$$

Therefore the eigenvalues of unitary matrix **A** are $\lambda_1 = \lambda_2 = i$ and $\lambda_3 = -i$.

Now $$\|\lambda_1\| = \sqrt{\langle \lambda_1, \lambda_1 \rangle} = \sqrt{\lambda_1 \bar{\lambda}_1}$$

$$= \sqrt{(i)(-i)}$$

$$= 1$$

Similarly $\|\lambda_2\| = \|\lambda_3\| = 1$. Therefore all the eigenvalues have the absolute value 1.

Normal Matrix

If **A** is a complex square matrix and

$$\mathbf{A}\bar{\mathbf{A}}^{\mathrm{T}} = \bar{\mathbf{A}}^{\mathrm{T}}\mathbf{A} \qquad \text{where } \bar{\mathbf{A}}^{\mathrm{T}} = \text{conjugate transpose of } \mathbf{A},$$

then **A** is called the *normal matrix*. For example, if

5.3 DIAGONALIZATION

In this section, we will discuss the diagonalization of a square matrix. We will also see that the eigenvectors play a central role in the diagonalization of a square matrix. In the first part of this section we will see that how to diagonalize the matrix and in the second part, we will discuss the problem of orthogonal diagonalization of a square matrix.

Let us start with the definition.

Definition: Diagonalizable Matrix

A square matrix **A** said to be *diagonalizable* if $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix for some invertible matrix **P** and the matrix **P** is said to diagonalize **A**.

EXAMPLE 5.30 Show that the matrix $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ is diagonalized to the matrix $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$.

Solution: The matrix **P** is said to be diagonalizable when $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix, that is, we have to show that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix.

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 5 & 3 \\ -5 & -6 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

So $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix. Hence the matrix **P** is diagonalized to matrix **A**.

The following theorem gives a characterization of diagonalizable matrix.

Theorem 5.16 [The Diagonalization Theorem]

An $n \times n$ matrix **A** is diagonalizable if and only if **A** has n linearly independent eigenvectors.

Here we are not taking the proof of the above theorem but its proof gives the following procedure for diagonalization.

Method to Diagonalize a Square Matrix A

Step 1 Find the linearly independent eigenvectors of **A**.

Step 2 Find the n linearly independent eigenvectors of **A**.

Step 3 Construct the matrix **P** having the n linearly independent eigenvectors as its column vectors.

Step 4 Construct a diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$ where diagonal entries are eigenvalues of **A**.

EXAMPLE 5.31 Diagonalize a matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$, if possible.

Solution: From Theorem 5.16, a 2×2 matrix **A** is diagonalizable if it has 2 linearly independent eigenvectors. So, first we find the eigenvectors of **A**.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}0 & 1\\-3 & 4\end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix}\lambda & -1\\3 & \lambda-4\end{bmatrix}\right) = 0$$

$$\lambda^2 - 4\lambda + 3 = 0 \qquad \text{(i)}$$

or $$(\lambda - 3)(\lambda - 1) = 0$$

$\therefore$ $$\lambda_1 = 3 \text{ or } \lambda_2 = 1$$

Step 2 ***Eigenvectors of* A**

***Case* (i)** For the eigenvalue $\lambda_1 = 3$:

Suppose $\mathbf{u} = [x_1, x_2]$ is an eigenvector of **A** corresponding to the eigenvalue $\lambda_1 = 3$. Then

$$(\lambda_1\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(3\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix}3 & -1\\3 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$3x_1 - x_2 = 0$$

$$x_1 = t,\ x_2 = 3t$$

$$\mathbf{u} = [x_1, x_2] = [t, 3t] = t[1, 3]$$

Thus $\mathbf{u} = \begin{bmatrix}1\\3\end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 3$.

***Case* (ii)** For the eigenvalue $\lambda_2 = 1$:

Suppose $\mathbf{v} = [y_1, y_2]$ is an eigenvector of **A** corresponding to the eigenvalue $\lambda_2 = 1$. Then

$$(\lambda_2\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix}1 & -1\\3 & -3\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$y_1 - y_2 = 0$$

$$y_1 = t,\ y_2 = t$$

$$\mathbf{v} = [y_1, y_2] = [t, t] = t[1, 1]$$

Thus $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 1$.

$$\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_2 = 0, x_1 = t \qquad \text{for any } t \in \mathbf{R}$$

$$\mathbf{u} = [x_1, x_2] = [t, 0] = t[1, 0]$$

Thus $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 5$. Therefore, we do not get two linearly independent eigenvalues for the 2×2 matrix **A**. Hence **A** is not a diagonalizable matrix.

EXAMPLE 5.33 Diagonalize the matrix $\mathbf{A} = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$, if possible.

Solution: A 3×3 matrix **A** is diagonalizable if it has 3 linearly independent eigenvectors. So, first we find the eigenvectors of **A**.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda+1 & -4 & 2 \\ 3 & \lambda-4 & 0 \\ 3 & -1 & \lambda-3 \end{bmatrix}\right) = 0$$

$$\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$$

or $$(\lambda - 1)(\lambda - 2)(\lambda - 3) = 0$$

$\therefore$ $$\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$$

Thus the 3×3 matrix **A** has 3 distinct eigenvalues, that is $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$. Hence **A** is diagonalizable.

Step 2 ***Eigenvectors of* A**

***Case* (i)** For the eigenvalue $\lambda_1 = 1$:

Suppose $\mathbf{u} = [x_1, x_2, x_3]$ is an eigenvector of **A** corresponding to the eigenvalue $\lambda_1 = 1$. Then

$$(\lambda_1\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix} 2 & -4 & 2 \\ 3 & -3 & 0 \\ 3 & -1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & -4 & 2 \\ 3 & -1 & 0 \\ 3 & -1 & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} -4 & 0 & 1 \\ 3 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$-4z_1 + z_3 = 0; \qquad 3z_1 - z_2 = 0$$

$$z_3 = 4z_1; \qquad z_2 = 3z_1$$

$$z_1 = t,\ z_2 = 3t,\ z_3 = 4t \qquad \text{for any } t \in R$$

$$\mathbf{w} = (z_1, z_2, z_3) = t[1, 3, 4]$$

Thus $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue $\lambda_3 = 3$.

Moreover, one can easily show that $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ are linearly independent eigenvectors.

Step 3 Take $\mathbf{P} = \begin{bmatrix} 1 & \frac{2}{3} & 1 \\ 1 & 1 & 3 \\ 1 & 1 & 4 \end{bmatrix}$. Then $\mathbf{P}^{-1} = \begin{bmatrix} 3 & -5 & 3 \\ -3 & 9 & -6 \\ 0 & -1 & 1 \end{bmatrix}$

Now $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 3 & -5 & 3 \\ -3 & 9 & -6 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & \frac{2}{3} & 1 \\ 1 & 1 & 3 \\ 1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

Therefore, the matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is the required diagonalized matrix of **A**.

EXAMPLE 5.34 Diagonalize the matrix $\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$, if possible.

Solution: A 3×3 matrix **A** is diagonalizable if it has 3 linearly independent eigenvectors. So, first we find the eigenvectors of **A**.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \mathbf{0}$$

$$\det\left(\lambda\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} - \begin{bmatrix}4 & 0 & 0\\1 & 4 & 0\\0 & 0 & 5\end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix}\lambda-4 & 0 & 0\\-1 & \lambda-4 & 0\\0 & 0 & \lambda-5\end{bmatrix}\right) = 0$$

or $$(\lambda-4)^2(\lambda-5) = 0$$

$\therefore$ $$\lambda_1 = \lambda_2 = 4, \quad \lambda_3 = 5$$

Thus the 3×3 matrix **A** has eigenvalues $\lambda_1 = \lambda_2 = 4$, $\lambda_3 = 5$.

Step 2 ***Eigenvectors of* A**

***Case* (i)** For the eigenvalue $\lambda_1 = \lambda_2 = 4$:

Suppose $\mathbf{u} = [x_1, x_2, x_3]$ is an eigenvector of **A** corresponding to eigenvalues $\lambda_1 = \lambda_2 = 4$. Then

$$(\lambda_1 \mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix}0 & 0 & 0\\-1 & 0 & 0\\0 & 0 & -5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$x_1 = x_3 = 0 \text{ and } x_2 = t \qquad \text{for any } t \in R$$

$$\mathbf{u} = [x_1, x_2, x_3] = t[0, 1, 0] \qquad \text{for any } t \in R$$

Thus $\mathbf{u} = \begin{bmatrix}0\\1\\0\end{bmatrix}$ is an eigenvector of **A** corresponding to eigenvalues $\lambda_1 = \lambda_2 = 4$.

***Case* (ii)** For the eigenvalue $\lambda_3 = 5$:

Suppose $\mathbf{v} = [y_1, y_2, y_3]$ is an eigenvector of **A** corresponding to the eigenvalue $\lambda_3 = 5$. Then

$$(\lambda_3 \mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(5\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Algebraic Multiplicity

The algebraic multiplicity of an eigenvalue λ of an $n \times n$ matrix **A** is the number of times λ occurs as a factor of the characteristic polynomial of matrix **A**.

Geometric Multiplicity

The geometric multiplicity of an eigenvalue λ of an $n \times n$ matrix **A** is the dimension of eigenspace of λ.

EXAMPLE 5.36 Find the eigenvalues and their algebraic and geometric multiplicity for the following matrices.

(i) $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ (iii) $\mathbf{C} = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

Solution:

(i) See Example 5.31, the characteristic equation of the given matrix **A** is

$$\lambda^2 - 4\lambda + 3 = 0 \qquad \text{(i)}$$

$$(\lambda - 3)(\lambda - 1) = 0$$

$$\lambda_1 = 3,\ \lambda_2 = 1$$

Therefore the algebraic multiplicity is 1 for both $\lambda_1 = 3$ and $\lambda_2 = 1$.
The eigenspaces are

$$E(3) = \left\{ t\begin{bmatrix} 1 \\ 3 \end{bmatrix} \middle| \ t \in R \right\}; \qquad \dim(E(3)) = 1$$

$$E(1) = \left\{ t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle| \ t \in R \right\}; \qquad \dim(E(1)) = 1$$

Therefore the geometric multiplicity is 1 for both $\lambda_1 = 3$ and $\lambda_2 = 1$.

(ii) See Example 5.32, the characteristic equation of the matrix **B** is

$$\lambda^2 - 10\lambda + 25 = 0$$

$$(\lambda - 5)^2 = 0$$

$$\lambda_1 = \lambda_2 = 5$$

Therefore the algebraic multiplicity of $\lambda = 5$ is 2.
The eigenspaces are

$$E(5)\left\{ t\begin{bmatrix} 1 \\ 0 \end{bmatrix} \middle| \ t \in R \right\}; \qquad \dim(E(5)) = 1$$

Therefore the geometric multiplicity of $\lambda = 5$ is 1.

(iii) See Example 5.4, the characteristic equation of the matrix **C** is

$$\lambda^3 - \lambda^2 - \lambda + 1 = 0 \qquad \text{(ii)}$$

$$(\lambda - 1)^2(\lambda + 1) = 0$$

$$\lambda = 1, -1$$

Therefore the algebraic multiplicity of $\lambda = 1$ is 2 and that of $\lambda = -1$ is 1.

The eigenspaces are

$$E(1) = \left\{ t\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \middle| \, t \in R \right\}; \qquad \dim(E(1)) = 2$$

$$E(-1) = \left\{ t\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} \middle| \, t \in R \right\}; \qquad \dim(E(-1)) = 1$$

Therefore the geometric multiplicity of $\lambda = 1$ is 2 and that of $\lambda = -1$ is 1.

From Example 5.36, we can conclude the following theorem.

Theorem 5.19 [Geometric Multiplicity]

Let **A** be an $n \times n$ matrix. Then the geometric multiplicity is less than or equal to the algebraic multiplicity for every eigenvalue of **A**.

The following theorem gives the condition for the diagonalization of an $n \times n$ matrix **A** that does not have n distinct eigenvalues.

Theorem 5.20 [Diagonalization of an $n \times n$ Matrix]

Let **A** be an $n \times n$ matrix. Then **A** is diagonalizable if and only if the geometric multiplicity is equal to the algebraic multiplicity for every eigenvalue of **A**.

EXAMPLE 5.37 Diagonalize the matrix $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 2 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}$, if possible.

Solution: A 4×4 matrix **A** is diagonalizable if its algebraic and geometric multiplicities are equal for each of its eigenvalues.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 2 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-2 & 0 & 0 & 0 \\ 0 & \lambda-2 & 0 & 0 \\ -3 & -2 & \lambda+1 & 0 \\ -1 & -1 & 0 & \lambda+1 \end{bmatrix}\right) = 0$$

or $$(\lambda + 1)^2 (\lambda - 2)^2 = 0$$

$\therefore$ $$\lambda = 2, -1$$

Therefore the algebraic multiplicity is 2 for both the eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -1$.

Step 2 ***Geometric multiplicity***

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = 2$.

Suppose $\mathbf{u} = [x_1, x_2, x_3, x_4] \in R^4$ is an eigenvector corresponding to eigenvalue $\lambda_1 = 2$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(2\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & -2 & 3 & 0 \\ -1 & -1 & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$3x_1 + 2x_2 - 3x_3 = 0; \qquad x_1 + x_2 - 3x_4 = 0$$

$$\begin{aligned} \mathbf{u} &= [x_1, x_2, x_3, x_4] \\ &= \left[t, s, t + \frac{2}{3}s, \frac{1}{3}t + \frac{1}{3}s\right] \\ &= t\left[1, 0, 1, \frac{1}{3}\right] + s\left[0, 1, \frac{2}{3}, \frac{1}{3}\right] \end{aligned}$$

Therefore $\begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{1}{3} \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ are eigenvectors of matrix $\mathbf{A}$ corresponding to the eigenvalue $\lambda_1 = 2$. Hence the eigenspace corresponding to the eigenvalue $\lambda_1 = 2$ is

$$E(2) = \left\{ t\begin{bmatrix} 1 \\ 0 \\ 1 \\ \frac{1}{3} \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \middle| \, t, s \in R \right\}; \qquad \dim(E(2)) = 2$$

So the geometric multiplicity of eigenvalue $\lambda_1 = 2$ is 2.

$$= [t, t]$$
$$= t[1, 1]$$

Thus, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$.

***Case* (ii)** For the eigenvalue $\lambda_2 = 2$.
Suppose $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = 1$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$
$$(2\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$
$$\begin{bmatrix} -2 & 3 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$2y_1 = 3y_2$$
$$y_1 = t, \quad y_2 = \frac{2}{3}t; \qquad \text{for any } t \in R$$
$$\mathbf{v} = [y_1, y_2]$$
$$= t\left[1, \frac{2}{3}\right]$$

Thus, $\mathbf{v} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.

Step 3 Construct $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & \frac{2}{3} \end{bmatrix}$. Then $\mathbf{P}^{-1} = \begin{bmatrix} -2 & 3 \\ 3 & -3 \end{bmatrix}$

And the diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

Step 4 From Theorem 5.21,

$$\mathbf{A}^k = \mathbf{PDP}^{-1}$$
$$\mathbf{A}^5 = \mathbf{PD}^5\mathbf{P}^{-1}$$
$$= \begin{bmatrix} 1 & 1 \\ 1 & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1^5 & 0 \\ 0 & 2^5 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 3 & -3 \end{bmatrix}$$
$$= \begin{bmatrix} 94 & -93 \\ 62 & -61 \end{bmatrix}$$

Solution: First, we will check the orthogonality of the matrix **P**.

$$\mathbf{P} = \begin{bmatrix} -\frac{4}{5} & 0 & \frac{3}{5} \\ \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \end{bmatrix}; \quad \mathbf{P}^{\mathrm{T}} = \begin{bmatrix} -\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \\ \frac{3}{5} & \frac{4}{5} & 0 \end{bmatrix}; \quad \mathbf{P}^{-1} = \begin{bmatrix} -\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \\ \frac{3}{5} & \frac{4}{5} & 0 \end{bmatrix} \quad \text{since } \det \mathbf{P} = -1$$

$$\mathbf{P}^{\mathrm{T}} = \mathbf{P}^{-1}$$

Therefore **P** is an orthogonal matrix.

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} \\ &= \begin{bmatrix} -\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \\ \frac{3}{5} & \frac{4}{5} & 0 \end{bmatrix} \begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -\frac{4}{5} & 0 & \frac{3}{5} \\ \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 25 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -50 \end{bmatrix} \quad \text{which is a diagonal matrix.} \end{aligned}$$

Therefore the matrix **P** is orthogonally diagonalized by the given matrix **A**.

The following theorem characterizes the orthogonally diagonalized real matrix.

Theorem 5.22 [Orthogonally Diagonalized Matrix]

An $n \times n$ real matrix **A** is orthogonally diagonalized if **A** is symmetric.

EXAMPLE 5.42 Which of the following matrices are orthogonally diagonalizable?

(i) $\mathbf{A} = \begin{bmatrix} 2 & -4 \\ 4 & 0 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$

Solution: By Theorem 5.22, a square matrix is orthogonally diagonalizable if it is a symmetric matrix. Therefore, we have to check the symmetricity of the given matrix.

(i) $\mathbf{A} = \begin{bmatrix} 2 & -4 \\ 4 & 0 \end{bmatrix} \quad \mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 2 & 4 \\ -4 & 0 \end{bmatrix}$

$$\mathbf{A} \neq \mathbf{A}^{\mathrm{T}}$$

Therefore, **A** is not a symmetric matrix. So **A** is not an orthogonally diagonalizable matrix.

(ii) $\mathbf{B} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \quad \mathbf{B}^{\mathrm{T}} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$

$$\mathbf{B} = \mathbf{B}^{\mathrm{T}}$$

Therefore, **B** is a symmetric matrix. So **B** is an orthogonally diagonalizable matrix.

EXAMPLE 5.43 Which of the following matrices are orthogonally diagonalizable?

(i) $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 1 \\ -3 & 0 & 1 \end{bmatrix}$ (ii) $\mathbf{D} = \begin{bmatrix} 3 & 0 & -4 \\ 0 & 1 & 2 \\ -4 & 2 & 1 \end{bmatrix}$

Solution: By Theorem 5.22, a square matrix is orthogonally diagonalizable if it is a symmetric matrix. Therefore, we have to check the symmetricity of the given matrix.

(i) $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 1 \\ -3 & 0 & 1 \end{bmatrix}$ $\mathbf{C}^{\mathrm{T}} = \begin{bmatrix} 1 & -2 & -3 \\ 2 & 0 & 0 \\ 3 & 1 & 1 \end{bmatrix}$

$\mathbf{C} \neq \mathbf{C}^{\mathrm{T}}$

Therefore **C** is not a symmetric matrix. So **C** is not an orthogonally diagonalizable matrix.

(ii) $\mathbf{D} = \begin{bmatrix} 3 & 0 & -4 \\ 0 & 1 & 2 \\ -4 & 2 & 1 \end{bmatrix}$ $\mathbf{D}^{\mathbf{T}} = \begin{bmatrix} 3 & 0 & -4 \\ 0 & 1 & 2 \\ -4 & 2 & 1 \end{bmatrix}$

$\mathbf{D} = \mathbf{D}^{\mathrm{T}}$

Therefore **D** is a symmetric matrix. So, **D** is an orthogonally diagonalizable matrix.

Using the following procedure, we can orthogonally diagonalize any square symmetric matrix.

Method to Orthogonalize a Square Symmetric Matrix

Let **A** be a symmetric matrix.

Step 1 Find the eigenvalues of the matrix **A**

Step 2 Find the basis of eigenspace of each eigenvalue of **A**.

Step 3 Construct the orthogonal basis for each eigenspace by applying the Gram-Schmidt process on the basis of eigenspace.

Step 4 Construct the matrix **P** having vectors of orthogonal bases as its column vectors.

Step 5 Calculate the diagonal matrix,

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}$$

EXAMPLE 5.44 Orthogonally diagonalize the matrix $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$.

Solution: Here **A** is a symmetric matrix. Therefore, **A** is an orthogonally diagonalizable matrix. We will use procedure enumerated above.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \mathbf{0}$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-1 & 2 \\ 2 & \lambda-1 \end{bmatrix}\right) = 0$$

or
$$\lambda^2 - 2\lambda - 3 = 0$$
$$(\lambda - 3)(\lambda + 1) = 0$$

$\therefore$
$$\lambda = 3 \text{ or } \lambda = -1$$

Therefore, $\lambda_1 = 3$ and $\lambda_2 = -1$ are the eigenvalues of matrix **A**.

Step 2 ***Eigenspaces of* A**

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = 3$.
Suppose $\mathbf{u} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda = 3$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{u} = 0$$
$$(3\mathbf{I} - \mathbf{A})\mathbf{u} = 0$$
$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$x_1 + x_2 = 0;$$
$$\mathbf{u} = [x_1, x_2)$$
$$= [x_1, -x_1]$$
$$= x_1[1, -1]$$

Thus, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 3$, and the eigenspace is

$$E(3) = \left\{ s\begin{bmatrix} 1 \\ -1 \end{bmatrix} \middle|\, s \in R \right\}$$

***Case* (ii)** Eigenspace for the eigenvalue $\lambda_2 = -1$
Suppose $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = -1$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$
$$(-\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$
$$\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$y_1 - y_2 = 0$$
$$y_1 = y_2 = t$$
$$\mathbf{v} = [y_1, y_2] = t[1, 1]$$

Thus, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = -1$, and the eigenspace is

$$E(-1) = \left\{ s\begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle|\, s \in R \right\}$$

$$\det\left(\begin{bmatrix} \lambda-3 & 2 & -4 \\ 2 & \lambda-6 & -2 \\ -4 & -2 & \lambda-3 \end{bmatrix}\right) = 0$$

$$\lambda^3 - 12\lambda^2 + 21\lambda + 98 = 0$$

or $$(\lambda - 7)^2 (\lambda + 2) = 0$$

$\therefore$ $$\lambda = 7, -2$$

Therefore $\lambda_1 = 7$ and $\lambda_2 = -2$ are the eigenvalues of the matrix **A**.

Step 2 ***Eigenspaces of* A**

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = 7$.

Suppose $\mathbf{u} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = 7$. Therefore ,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(7\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 2 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$2x_1 + x_2 - 2x_3 = 0$$

$$x_1 = s,\ x_2 = 2t - 2\mathrm{s},\ x_3 = t$$

$$\begin{aligned} \mathbf{u} &= [x_1, x_2, x_3] \\ &= [s, 2t - 2s, t] \\ &= s[1, -2, 0] + t[0, 2, 1] \end{aligned}$$

Therefore $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ are the eigenvectors of **A** corresponding to the eigenvalue $\lambda_1 = 7$.

$$\text{Eigenspace } E(7) = \left\{ s\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \middle| s, t \in R \right\}$$

***Case* (ii)** Eigenspace for the eigenvalue $\lambda_2 = -2$.

Suppose $\mathbf{v} = [y_1, y_2, y_3] \in R^3$ is an eigenvector corresponding to the eigenvalue $\lambda = -2$. Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(-2\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix} -5 & 2 & -4 \\ 2 & -8 & -2 \\ -4 & -2 & -5 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$y_1 = 2y_2 \quad \text{and} \quad y_3 = -2y_2$$

$$y_1 = 2t, \; y_2 = t, \; y_3 = -2t$$

$$\begin{aligned} \mathbf{u} &= [x_1, x_2, x_3] \\ &= [2t, t, -2t] \\ &= t[2, 1, -2] \end{aligned}$$

Therefore, $\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ is the eigenvector of **A** corresponding to the eigenvalue $\lambda_2 = -2$ and

$$\text{Eigenspace } E(-2) = \left\{ t\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \middle| \, t \in R \right\}$$

Step 3 ***Orthonormal basis: Gram–Schmidt process***

By applying the Gram–Schmidt process to the eigenspace $E(7)$, we get

$$\mathbf{v}_1 = \mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \\ 0 \end{bmatrix}$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \\ 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} \frac{4}{3\sqrt{5}} \\ \frac{2}{3\sqrt{5}} \\ \frac{\sqrt{5}}{3} \end{bmatrix}$$

Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthogonormal basis for the eigenvalue $E(7)$.

Similarly $\mathbf{w}_3 = \dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} \dfrac{2}{3} \\ \dfrac{1}{3} \\ -\dfrac{2}{3} \end{bmatrix}$ $\{\mathbf{w}_3\}$ is an orthonormal basis for the eigenspace $E(-2)$.

Step 4 ***Construct the matrix* P**

$$\text{Take } \mathbf{P} = [\mathbf{w}_1\ \mathbf{w}_2\ \mathbf{w}_3] = \begin{bmatrix} \dfrac{1}{\sqrt{5}} & \dfrac{4}{3\sqrt{5}} & \dfrac{2}{3} \\ -\dfrac{2}{\sqrt{5}} & \dfrac{2}{3\sqrt{5}} & \dfrac{1}{3} \\ 0 & \dfrac{\sqrt{5}}{3} & -\dfrac{2}{3} \end{bmatrix}$$

It is left to the reader to prove the orthogonality of **P**, that is $\mathbf{P}^{-1} = P^{\mathrm{T}}$.

Step 5 ***Calculate the diagonal matrix* D**

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Unitary Diagonalization

Here we will consider the complex square matrix that is the matrix with complex entries. Let us start with a definition.

Definition: Unitary Diagonalization

A complex square matrix **A** is called unitary diagonalizable if there exists a unitary matrix **P** such that $\mathbf{P}^{-1}\mathbf{AP}$ (= $\mathbf{P}^*\mathbf{AP}$) is diagonal and **P** is then said to unitarily diagonalize **A**.

Remark: An $n \times n$ complex matrix **A** is *unitary* if and only if the row vectors (or column vectors) of **A** form an orthonormal set in C^n.

EXAMPLE 5.46 Is the matrix $\mathbf{A} = \begin{bmatrix} 6 & 2+2i \\ 2-2i & 4 \end{bmatrix}$ unitarily diagonalized by the matrix

$$\mathbf{P} = \begin{bmatrix} -\dfrac{1+i}{\sqrt{6}} & \dfrac{1+i}{\sqrt{3}} \\ \dfrac{2}{\sqrt{6}} & \dfrac{1}{\sqrt{3}} \end{bmatrix}?$$

Solution: For the given matrix,

$$\mathbf{P} = \begin{bmatrix} -\frac{1+i}{\sqrt{6}} & \frac{1+i}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\mathbf{P}^* = \overline{\mathbf{P}}^{\mathrm{T}} = \begin{bmatrix} -\frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1-i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \quad \text{and} \quad \mathbf{P}^{-1} = \begin{bmatrix} -\frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1-i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\mathbf{P}^* = \mathbf{P}^{-1}$$

Therefore, P is a unitary matrix.

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^*\mathbf{AP} = \begin{bmatrix} -\frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1-i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 6 & 2+2i \\ 2-2i & 4 \end{bmatrix} \begin{bmatrix} -\frac{1+i}{\sqrt{6}} & \frac{1+i}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$$

Hence the given matrix **P** is unitarily diagonalized by the given matrix **A**.

Theorem 5.23 [Unitary Diagonalization]

An $n \times n$ complex matrix **A** is unitarily diagonalized if **A** is normal.

EXAMPLE 5.47 Which of the following matrices is unitary diagonalizable?

(i) $\mathbf{A} = \begin{bmatrix} i & i \\ i & -i \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 1+i & 1-i \\ \sqrt{2} & -\sqrt{2} \end{bmatrix}$

Solution:

(i) For $\mathbf{A} = \begin{bmatrix} i & i \\ i & -i \end{bmatrix}$

$$\mathbf{A}^* = \overline{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} -i & -i \\ -i & i \end{bmatrix}$$

$$\mathbf{AA}^* = \begin{bmatrix} i & i \\ i & -i \end{bmatrix}\begin{bmatrix} -i & -i \\ -i & i \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{A}^*\mathbf{A} = \begin{bmatrix} -i & -i \\ -i & i \end{bmatrix}\begin{bmatrix} i & i \\ i & -i \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$$

So **A** is a normal matrix. Therefore, A is unitary diagonalizable.

(i) For $\mathbf{B} = \begin{bmatrix} 1+i & 1-i \\ \sqrt{2} & -\sqrt{2} \end{bmatrix}$

$$\mathbf{B}^* = \begin{bmatrix} 1-i & \sqrt{2} \\ 1+i & -\sqrt{2} \end{bmatrix}$$

$$\mathbf{BB}^* = \begin{bmatrix} 1+i & 1-i \\ \sqrt{2} & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 1-i & \sqrt{2} \\ 1+i & -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 4 & 2\sqrt{2}i \\ -2\sqrt{2}i & 4 \end{bmatrix}$$

$$\mathbf{B}^*\mathbf{B} = \begin{bmatrix} 1-i & \sqrt{2} \\ 1+i & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 1+i & 1-i \\ \sqrt{2} & -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 4 & -2i-2 \\ 2i-2 & 4 \end{bmatrix}$$

$$\mathbf{BB}^* \neq \mathbf{B}^*\mathbf{B}$$

So **B** is not a normal matrix. Therefore, **B** is not unitary diagonalizable.

Remark: The procedure to unitarily diagonalize a complex square matrix is exactly the same as the procedure to orthogonally diagonalize a real square matrix.

EXAMPLE 5.48 Unitary diagonalize a matrix $\mathbf{A} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}$.

Solution: Here **A** is a normal matrix. So **A** is a unitary diagonalizable matrix.

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda-2 & -3+3i \\ -3-3i & \lambda-5 \end{bmatrix}\right) = 0$$

$$\lambda^2 - 7\lambda - 8 = 0$$

or $$(\lambda - 8)(\lambda + 1) = 0$$

∴ $$\lambda = -1 \text{ or } \lambda = 8$$

Therefore $\lambda_1 = -1$ and $\lambda_2 = 8$ are the eigenvalues of matrix **A**

Step 2 ***Eigenspaces of* A**

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = -1$.

Suppose $\mathbf{u} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = -1$.

Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(-\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix} -3 & -3+3i \\ -3-3i & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 1 & 1-i \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 + (1 - i)x_2 = 0$$

$$x_2 = t, \; x_1 = (-1 + i)t$$

$$\mathbf{u} = [x_1, x_2]$$

$$= t[-1 + i, 1]$$

Thus, $\begin{bmatrix} -1+i \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = -1$, and the eigenspace

$$\mathbf{E}(-1) = \left\{ s\begin{bmatrix} -1+i \\ 1 \end{bmatrix} \middle| \, s \in R \right\}$$

That is, $\begin{bmatrix} -1+i \\ 1 \end{bmatrix}$ is a basis vector for the eigenspace $E(-1)$.

***Case* (ii)** Eigenspace for the eigenvalue $\lambda_2 = 8$.
Suppose $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 8$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(8\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix} 6 & -3+3i \\ -3-3i & 3 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 2 & -1+i \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$2y_1 + (-1 + i)\, y_2 = 0;$$

$$y_2 = (1 + i)\, y_1$$

$$\mathbf{v} = [y_1, y_2] = y_1[1, 1 + i]$$

Thus, $\begin{bmatrix} 1 \\ 1+i \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 8$, and the eigenspace

$$E(8) = \left\{ s\begin{bmatrix} 1 \\ 1+i \end{bmatrix} \middle| \, s \in R \right\}$$

That is, $\begin{bmatrix} 1 \\ 1+i \end{bmatrix}$ is a basis vector for the eigenspace $E(8)$.

$$= \lambda_1 w_1 w_1^{\mathrm{T}} + \lambda_2 w_2 w_2^{\mathrm{T}} + \cdots + \lambda_n w_n w_n^{\mathrm{T}}$$

This representation of **A** is called the *spectral decomposition* of **A**.

EXAMPLE 5.49 Construct a spectral decomposition of $\mathbf{A} = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

Solution: The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 1 & -5 \\ -5 & \lambda - 1 \end{bmatrix}\right) = 0$$

or $$\lambda^2 - 2\lambda - 24 = 0$$

$$(\lambda + 4)(\lambda - 6) = 0$$

$\therefore$ $$\lambda = -4 \quad \text{or} \quad \lambda = 6$$

Hence $\lambda_1 = -4$ and $\lambda_2 = 6$ are the eigenvalues of the matrix **A**

Step 2 ***Eigenspaces of* A**

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = -4$.
Suppose $\mathbf{u} = (x_1, x_2) \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = -4$.
Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{u} = 0$$

$$(-4\mathbf{I} - \mathbf{A})\mathbf{u} = 0$$

$$\begin{bmatrix} -5 & -5 \\ -5 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 + x_2 = 0$$

$$x_2 = -x_1$$

$$\mathbf{u} = [x_1, -x_1]$$

$$= x_1[1, -1]$$

Thus $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = -4$, and the eigenspace

$$E(-4) = \left\{ s\begin{bmatrix} 1 \\ -1 \end{bmatrix} \middle| \, s \in R \right\}$$

That is, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is a basis vector for the eigenspace $E(-4)$.

***Case* (ii)** Eigenspace for the eigenvalue $\lambda_2 = 6$.
Suppose $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 6$.
Therefore ,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(6\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$\begin{bmatrix} 5 & -5 \\ -5 & 5 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$y_1 - y_2 = 0;$$

$$y_2 = y_1$$

$$\mathbf{v} = [y_1, y_2] = y_1[1, 1]$$

Thus, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 6$, and the eigenspace

$$E(6) = \left\{ s\begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle| \, s \in R \right\}$$

That is, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a basis vector for the eigenspace $E(6)$.

Step 3 ***Orthogonal basis: Gram–Schmidt process***
By applying the Gram–Schmidt process on eigenspaces $E(-4)$ and $E(6)$, we get

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{w}_1 = \frac{\mathbf{u}}{\|\mathbf{u}\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{w}_2 = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Therefore the spectral decomposition of the matrix $\mathbf{A}$ is

$$\mathbf{A} = \lambda_1 \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \lambda_2 \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}}$$

$$= -4\mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + 6\mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} \qquad \text{where } \mathbf{w}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}; \ \mathbf{w}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

EXAMPLE 5.50 Construct the spectral decomposition of $\mathbf{A} = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$.

Solution: The characteristic equation of **A** is

Step 1 ***Eigenvalues of* A**

The characteristic equation of the matrix **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 5 & 4 & 2 \\ 4 & \lambda - 5 & -2 \\ 2 & -2 & \lambda - 2 \end{bmatrix}\right) = 0$$

or $$\lambda^3 - 12\lambda^2 + 21\lambda - 10 = 0$$

$$(\lambda - 1)^2 (\lambda - 10) = 0$$

$\therefore$ $$\lambda_1 = 1,\ \lambda_2 = 1,\ \lambda_3 = -10$$

Therefore $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = -10$ are the eigenvalues of the matrix **A**.

Step 2 ***Eigenspaces of* A**

***Case* (i)** Eigenspace for the eigenvalue $\lambda_1 = \lambda_2 = 1$.

Suppose $\mathbf{u} = [x_1, x_2, x_3] \in R^3$ is an eigenvector corresponding to the eigenvalue 1.

Therefore,

$$(\lambda \mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$(\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

$$\begin{bmatrix} -4 & 4 & 2 \\ 4 & -4 & -2 \\ 2 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\sim \quad \begin{bmatrix} 2 & -2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$2x_1 - 2x_2 - x_3 = 0$$

$$x_1 = t,\ x_2 = s,\ x_3 = 2t - 2s,$$

$$\mathbf{u} = [x_1, x_2, x_3]$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} \frac{4}{5} \\ 1 \\ -\frac{2}{5} \end{bmatrix} \quad \Rightarrow \quad \mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} \frac{4}{3\sqrt{5}} \\ \frac{\sqrt{5}}{3} \\ -\frac{2}{3\sqrt{5}} \end{bmatrix}$$

Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthonormal basis for the eigenvalue $E(1)$.

Similarly $\mathbf{w}_3 = \dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ $\{\mathbf{w}_3\}$ is an orthonormal basis for the eigenspace $E(-10)$.

Therefore, the spectral decomposition of the matrix **A** is

$$\begin{aligned} \mathbf{A} &= \lambda_1 \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \lambda_2 \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} + \lambda_3 \mathbf{w}_3 \mathbf{w}_3^{\mathrm{T}} \\ &= \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} + 10 \mathbf{w}_3 \mathbf{w}_3^{\mathrm{T}} \end{aligned}$$

where

$$\mathbf{w}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} \frac{4}{3\sqrt{5}} \\ \frac{\sqrt{5}}{3} \\ -\frac{2}{3\sqrt{5}} \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

EXERCISE SET 4

1. Is $\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ orthogonally diagonalized by $\mathbf{P} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$?

2. Orthogonally diagonalize the following matrices:

(i) $\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$

3. Unitarily diagonalize the matrices given below:

(i) $\begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & -1+i \\ 0 & -1-i & 0 \end{bmatrix}$

4. Construct the spectral decomposition of the following matrices:

(i) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}$ (iii) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

5.5 APPLICATION TO QUADRATIC FORMS

In this section, we will see the applications of eigenvalues and eigenvectors to the quadratic forms. First we will define the quadratic form and its matrix representation. Then we will discuss some problems on quadratic forms: maximum and minimum values of quadratic forms with some constraints, positive definite form of quadratic forms, diagonalization of quadratic forms and the principal axes theorems.

Let us begin with the definition.

Definition: Quadratic Form

The quadratic form in n variables $x_1, x_2, x_3, \ldots, x_n$ is

$$\begin{aligned} & a_{11}x_1^2 + a_{12}\,x_1x_2 + \cdots + a_{1n}\,x_1x_n \\ & + a_{21}\,x_2x_1 + a_{22}\,x_2^2 + \cdots + a_{2n}\,x_2x_n \\ & + \quad \vdots \qquad \vdots \qquad \cdots \qquad \vdots \\ & + a_{n1}\,x_nx_1 + a_{n2}\,x_nx_2 + \cdots + a_{nn}\,x_n^2 \end{aligned}$$

where all the coefficient $a_{ij}(1 \le i \le j \le n)$ are real scalars.

Note that the terms of the form $a_{ij}x_ix_j$ $(i \ne j)$ in a quadratic form are called *cross product terms*.

Moreover the quadratic form can be expressed as a matrix multiplication,

$$\begin{aligned} & a_{11}x_1^2 + a_{12}\,x_1x_2 + \cdots + a_{1n}\,x_1x_n + a_{21}\,x_2x_1 + a_{22}\,x_2^2 + \cdots + a_{2n}\,x_2x_n + \cdots + a_{n1}\,x_nx_1 \\ & + a_{n2}\,x_nx_2 + \cdots + a_{nn}\,x_n^2 \\ & = [x_1\ x_2 \ldots x_n]\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ . & . & \ldots & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ & = \mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ . & . & \ldots & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$ and $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

EXAMPLE 5.51 Generate the quadratic forms from the following matrices.

(i) $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 4 & -3 \end{bmatrix}$ (ii) $\mathbf{B} = \begin{bmatrix} 1 & 5 \\ 3 & -3 \end{bmatrix}$

(iii) $\mathbf{C} = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 2 & 5 \\ -2 & 5 & 3 \end{bmatrix}$ (iv) $\mathbf{D} = \begin{bmatrix} 1 & 3 & -\frac{3}{2} \\ 1 & 2 & 6 \\ -\frac{5}{2} & 4 & 3 \end{bmatrix}$

Solution:

(i) The quadratic form for the given matrix $\mathbf{A}_1$ is

$$\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} = [x_1\ x_2]\begin{bmatrix} 1 & 4 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= x_1^2 + 8x_1x_2 - 3x_2^2$$

(ii) The quadratic form for the given matrix **B** is

$$\mathbf{X}^{\mathrm{T}}\mathbf{B}\mathbf{X} = [x_1\ x_2]\begin{bmatrix} 1 & 5 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= x_1^2 + 8x_1x_2 - 3x_2^2$$

(iii) The quadratic form for the given matrix **C** is

$$\mathbf{X}^{\mathrm{T}}\mathbf{C}\mathbf{X} = [x_1\ x_2\ x_3]\begin{bmatrix} 1 & 2 & -2 \\ 2 & 2 & 5 \\ -2 & 5 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 10x_2x_3 - 4x_1x_3$$

(iv) The quadratic form for the given matrix **D** is

$$\mathbf{X}^{\mathrm{T}}\mathbf{A}_4\mathbf{X} = [x_1\ x_2\ x_3]\begin{bmatrix} 1 & 3 & -\frac{3}{2} \\ 1 & 2 & 6 \\ -\frac{5}{2} & 4 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 10x_2x_3 - 4x_1x_3$$

From the above example, we can conclude that the different matrices can be expressed by the same quadratic form. Also note that, in the above example, **A** and **C** are symmetric while **B** and **D** are not symmetric. Since we are more comfortable to work with symmetric matrix, so here without any restriction we consider only symmetric matrices to represent the quadratic form.

Therefore, if $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is a matrix representation of the quadratic form, then **A** is a symmetric matrix whose diagonal entries are the coefficients of the square terms and the off-diagonal entries are half the coefficients of the cross product terms.

Theorem 5.25 [Symmetric Matrix representation of quadratic forms]

For any $n \times n$ matrix **A** and any $1 \times n$ row matrix **X**, we have $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \mathbf{X}^{\mathrm{T}}\mathbf{BX}$ where $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$ is the symmetric matrix.

EXAMPLE 5.52 Find the matrix representation of the following quadratic forms:

(i) $x_1^2 - 2x_1x_2 + 3x_2^2$ (ii) $x_1^2 + 3x_2^2 - 10x_3^2 + x_1x_2 - 8x_1x_3$

(iii) $x_1^2 + 2x_2^2$ (iv) $x_1^2 - x_2^2 + 4x_3^2$

Solution:

(i) $x_1^2 + 3x_2^2 - 2x_1x_2 = [x_1\ x_2]\begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

(ii) $x_1^2 + 3x_2^2 - 10x_3^2 + x_1x_2 - 8x_1x_3 = [x_1\ x_2\ x_3]\begin{bmatrix} 1 & \frac{1}{2} & -4 \\ \frac{1}{2} & 3 & 0 \\ -4 & 0 & -10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(iii) $x_1^2 + 2x_2^2 = [x_1\ x_2]\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

(iv) $x_1^2 - x_2^2 + 4x_3^2 = [x_1\ x_2\ x_3]\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Classification of Quadratic Forms

The classification of quadratic forms depends on the values of quadratic form as given below.

Definition: Classification of Quadratic Forms

Let **A** be a symmetric matrix. The quadratic form is $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ called

(a) positive definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX} > 0$ for all $x \neq 0$
(b) semi-positive definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX} \geq 0$ for all x
(c) negative definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX} < 0$ for all $x \neq 0$
(d) semi-negative definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX} \leq 0$ for all x
(e) indefinite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ has both positive and negative values

We will have a detailed discussion on positive definite quadratic form. Note that a symmetric matrix **A** is called positive definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is positive definite.

Positive Definite Forms

Here we will see two different methods to determine whether a quadratic form is positive definite or not. The first method will be used with eigenvalues while the second will be explained without using eigenvalues.

The following theorem characterizes the positive definite quadratic form in terms of eigenvalues.

Theorem 5.26 [Characterization of Positive Definite Form]

Let **A** be a symmetric matrix. Then the following statements are equivalent.

(a) The quadratic form $\mathbf{X}^T\mathbf{A}\mathbf{X}$ is positive definite.
(b) The matrix **A** is positive definite.
(c) All the eigenvalues of **A** are positive.

EXAMPLE 5.53 Which of the following quadratic forms are positive definite?

(i) $5x_1^2 - 2x_1x_2 + 5x_2^2$

(ii) $2x_1^2 - 6x_1x_2 - 6x_2^2$

(iii) $6x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 2x_1x_3 - 2x_2x_3$

(iv) $3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$

Solution:

(i) The matrix representation of the given quadratic form is

$$5x_1^2 - 2x_1x_2 + 5x_2^2 = [x_1\ x_2]\begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Take $\mathbf{A} = \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$. By Theorem 5.26, the given quadratic form is positive definite if all the eigenvalues of **A** are positive.

The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 5 & 1 \\ 1 & \lambda - 5 \end{bmatrix}\right) = 0$$

or $$\lambda^2 - 10\lambda + 24 = 0$$

or $$(\lambda - 6)(\lambda - 4) = 0$$

$\therefore$ $$\lambda = 6 \text{ or } \lambda = 4$$

Thus, the eigenvalues 6 and 4 are positive. Therefore, the given quadratic form is positive definite.

(ii) The matrix representation of the given quadratic form is

$$2x_1^2 - 6x_1x_2 - 6x_2^2 = [x_1\ x_2]\begin{bmatrix} 2 & -3 \\ -3 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Take $\mathbf{A} = \begin{bmatrix} 2 & -3 \\ -3 & -6 \end{bmatrix}$. By Theorem 5.26, the given quadratic form is positive definite if all the eigenvalues of **A** are positive.

The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & -3 \\ -3 & -6 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 2 & 3 \\ 3 & \lambda + 6 \end{bmatrix}\right) = 0$$

or $$\lambda^2 + 4\lambda - 21 = 0$$

or $$(\lambda - 3)(\lambda + 7) = 0$$

$\therefore$ $$\lambda = 3 \text{ or } \lambda = -7$$

Thus, all the eigenvalues are not positive. Therefore, the given quadratic form is not positive definite.

(iii) The matrix representation of the given quadratic form is

$$6x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 2x_1x_3 - 2x_2x_3 = [x_1\ x_2\ x_3]\begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Take $\mathbf{A} = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$. By Theorem 5.26, the given quadratic form is positive definite if all the eigenvalues of **A** are positive.

The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 6 & 2 & 1 \\ 2 & \lambda - 6 & 1 \\ 1 & 1 & \lambda - 5 \end{bmatrix}\right) = 0$$

or $$\lambda^3 - 17\lambda^2 + 90\lambda - 144 = 0$$

or $$(\lambda - 8)(\lambda - 6)(\lambda - 3) = 0$$

∴ $$\lambda = 8 \text{ or } \lambda = 6 \text{ or } \lambda = 3$$

Thus, all the eigenvalues 8, 6 and 3 are positive. Therefore, the given quadratic form is positive definite.

(iv) The matrix representation of the given quadratic form is

$$3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3 = [x_1\ x_2\ x_3]\begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Take $\mathbf{A} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. By Theorem 5.26, the given quadratic form is positive definite if all the eigenvalues of **A** are positive.

The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 3 & -2 & 0 \\ -2 & \lambda - 2 & -2 \\ 0 & -2 & \lambda - 1 \end{bmatrix}\right)$$

or $$\lambda^3 - 6\lambda^2 + 3\lambda + 10 = 0$$

or $$(\lambda - 5)(\lambda - 2)(\lambda + 1) = 0$$

∴ $$\lambda = 5 \text{ or } \lambda = 2 \text{ or } \lambda = -1$$

Thus, all the eigenvalues 5, 2 and – 1 are all not positive. Therefore the given quadratic form is not positive definite.

Now we will discuss the second method for the positive definite quadratic form.

Definition: Principal Submatrix

Let **A** be an $n \times n$ square matrix. Then the *principal submatrix* $\mathbf{A}_r$ of order r of **A** is the submatrix formed from the first r rows and r columns of **A** for $r = 1, 2, \ldots, n$.

If $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, then the principal submatrices are

The determinants of the principal submatrices are

$$|\mathbf{A}_1| = |6| = 6 \qquad |\mathbf{A}_2| = \begin{vmatrix} 6 & -2 \\ -2 & 6 \end{vmatrix} = 32 \qquad |\mathbf{A}_3| = \begin{vmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{vmatrix} = 144$$

Here the determinant of every principal submatrix **A** is positive definite.

EXAMPLE 5.55 For which values of k, the quadratic form $x_1^2 - 6x_1x_2 + kx_2^2$ is positive?

Solution: The matrix form of the given quadratic form is

$$x_1^2 - 6x_1x_2 + kx_2^2 = [x_1 \ x_2]\begin{bmatrix} 1 & -3 \\ -3 & k \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

By Theorem 5.27, the determinant of every principal submatrix is positive. The determinants of the principal submatrices are

$$|\mathbf{A}_1| = |1| = 1 \qquad |\mathbf{A}_2| = \begin{vmatrix} 1 & -3 \\ -3 & k \end{vmatrix} = k - 9$$

Thus the given quadratic form is positive definite if

$$|\mathbf{A}_2| > 0 \qquad \Rightarrow \qquad k - 9 > 0 \qquad \Rightarrow \qquad k > 9$$

Therefore for all k greater than 9 the given quadratic form is positive definite.

Maximum and Minimum Values of the Quadratic Forms

Here we will calculate the maximum and minimum values of the quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ subject to the constraint

$$||\mathbf{X}|| = |\mathbf{X}^{\mathrm{T}}\mathbf{X} = \sqrt{(x_1^2 + x_2^2 + \cdots + x_n^2)}$$

The following theorem gives the solution of the above problem.

Theorem 5.28 [Maximum and Minimum Values of the Quadratic Form]

Let **A** be a symmetric $n \times n$ matrix. If the eigenvalues of **A** are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and $||\mathbf{X}|| = 1$, then

(i) $\lambda_1 \geq \mathbf{X}^{\mathrm{T}}\mathbf{AX} \geq \lambda_n$

(ii) $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \lambda_1$ if **X** is an eigenvector corresponding to λ_1 and $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \lambda_n$ if **X** is an eigenvector corresponding to λ_n.

This theorem shows that the largest eigenvalue of **A** is the maximum value of the quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ and the smallest eigenvalue of **A** is the minimum value of the quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ subject to the constraint $||\mathbf{X}|| = 1$.

Procedure to Find the Maximum and Minimum Values of Quadratic Form

Let **A** be a symmetric matrix.

Step 1 Find the eigenvalues of **A**.

Step 2 The maximum value of the quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is equal to the largest eigenvalue of **A**. The minimum value of the quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is equal to the smallest eigenvalue of **A**.

Thus, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 7$. So the vector at which the given quadratic form attained maximum value is $\mathbf{w}_1 = \dfrac{\mathbf{u}}{\|\mathbf{u}\|} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} \end{bmatrix}$

***Case* (ii)** Suppose the vector $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to an eigenvalue $\lambda_1 = 3$. Therefore,

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0$$

$$(3\mathbf{I} - \mathbf{A})\mathbf{v} = 0$$

$$\begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$y_1 = y_2$$

$$\begin{aligned} \mathbf{v} &= [y_1, y_2] \\ &= [y_1, y_1] \\ &= y_1[1, 1] \end{aligned}$$

Thus, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_3 = 3$. So the vector at which a given quadratic form attained minimum value is $\mathbf{w}_2 = \dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}$

Diagonalization of Quadratic Form

The diagonalization of a quadratic form is a process of removing the cross product terms from a quadratic form. In other words, it is the process to reduce the quadratic form into a sum of squares. To do this, we will use the change of variables method as given below.

Let $\mathbf{A}$ be an $n \times n$ symmetric matrix and $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}$ be a quadratic form. If we change the variable x into y by substituting

$$\mathbf{X} = \mathbf{P}\mathbf{Y}$$

where the matrix $\mathbf{P}$ is orthogonally diagonalized to a matrix $\mathbf{A}$, that is, if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of $\mathbf{A}$ then

$$\mathbf{D} = \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} 3$$

Hence, we get

$$\begin{aligned}\mathbf{X}^{\mathrm{T}}\mathbf{AX} &= (\mathbf{PY})^{\mathrm{T}}\mathbf{A}(\mathbf{PY})\\ &= \mathbf{Y}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{APY}\\ &= \mathbf{Y}^{\mathrm{T}}\mathbf{DY}\\ &= [y_1\ y_2 \cdots y_n]\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ . & . & \cdots & .\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}\\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2\end{aligned}$$

Hence the quadratic form is represented as a sum of squares or with no cross product term. From the above discussion, we have the following theorem.

Theorem 5.29 [Diagonalization of Quadratic Forms]

Let **A** be an $n \times n$ symmetric matrix. If there is a change of variable by substituting $\mathbf{X} = \mathbf{PY}$, then it transforms the quadratic $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ into a quadratic of the form $\mathbf{Y}^{\mathrm{T}}\mathbf{AY}$ with no cross product terms.

EXAMPLE 5.57 Find the change of variable that transforms the quadratic form $x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross product terms.

Solution: The given quadratic form can be represented in matrix form as

$$x_1^2 - 8x_1x_2 - 5x_2^2 = [x_1\ x_2]\begin{bmatrix}1 & -4\\ -4 & -5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

Step 1 Take $\mathbf{A} = \begin{bmatrix}1 & -4\\ -4 & -5\end{bmatrix}$. The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - \begin{bmatrix}1 & -4\\ -4 & -5\end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix}\lambda - 1 & 4\\ 4 & \lambda + 5\end{bmatrix}\right) = 0$$

or $$\lambda^2 + 4\lambda - 21 = 0$$

or $$(\lambda - 3)(\lambda + 7) = 0$$

$\therefore$ $$\lambda = 3 \text{ or } \lambda = -7$$

Thus $\lambda_1 = 3$ and $\lambda_2 = -7$ are the eigenvalues of **A**. One can easily calculate the unit eigenvectors

$\mathbf{w}_1 = \begin{bmatrix}\dfrac{2}{\sqrt{5}}\\ -\dfrac{1}{\sqrt{5}}\end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix}\dfrac{1}{\sqrt{5}}\\ \dfrac{2}{\sqrt{5}}\end{bmatrix}$ corresponding to eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -7$ respectively.

Construct a matrix $\mathbf{P} = [\mathbf{w}_1\ \mathbf{w}_2] = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$

Diagonal matrix $\mathbf{D} = \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$

By substituting $\mathbf{X} = \mathbf{P}\mathbf{Y}$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

i.e. $$\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} = \mathbf{Y}^{\mathrm{T}}\mathbf{D}\mathbf{Y} = 3y_1^2 - 7y_2^2$$

This is the required quadratic form with no cross product terms.

Principal Axes Theorem

Here we will discuss the principal axes theorem in two ways: one for the plane curves in R^2 and the second for the surfaces in R^3. These theorems are simply generalizations of Theorem 5.29 for the conic sections. We will also discuss their geometrical interpretations.

***Quadratic Curves in** R^2*

We consider here the quadratic equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \tag{5.1}$$

where a, b, c, d, e and f are real numbers and a, b and c are not all zero. In the quadratic Eq. (5.1), the term

$$ax^2 + 2bxy + cy^2$$

is called the *associated quadratic form*.

Theorem 5.30 [Principal Axes Theorem in R^2]

If the equation of the conic in the xy-co-ordinate system is

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

and the associated quadratic form is

$$\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} = ax^2 + 2bxy + cy^2$$

then the co-ordinate axes of the xy-co-ordinate system can be rotated on to the co-ordinate axes of the $x'y'$-co-ordinate system such that the equation of the conic in the new $x'y'$-co-ordinate system is

$$\lambda_1 x'^2 + 2bx'y' + \lambda_2 y'^2 + d'x' + e'y' + f = 0$$

where λ_1 and λ_2 are the eigenvalues of $\mathbf{A}$. Moreover, the above rotation can be done by the change of variable method $\mathbf{X} = \mathbf{P}\mathbf{X}'$ where $\mathbf{P}$ orthogonally diagonalizes $\mathbf{A}$ and det $\mathbf{P} = 1$.

Thus, $\lambda_1 = 3$ and $\lambda_2 = -2$ are the eigenvalues of **A**. Therefore $\lambda_1 = 3$ and $\lambda_2 = -2$ are the eigenvalues of **A**. It is easy to find that $\mathbf{w}_1 = \begin{bmatrix} -\dfrac{2}{\sqrt{5}} \\ \dfrac{1}{\sqrt{5}} \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} \dfrac{1}{\sqrt{5}} \\ \dfrac{2}{\sqrt{5}} \end{bmatrix}$ are orthogonal bases for the eigenspaces corresponding to eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -2$ respectively.

So, we can construct the matrix

$$\mathbf{P} = [\mathbf{w}_2 \ \mathbf{w}_1] = \begin{bmatrix} \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} \\ \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \end{bmatrix}$$

Rotate the co-ordinate axes of the xy-co-ordinate system by substituting

$$\mathbf{X} = \mathbf{PX}'$$

i.e.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} \\ \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

in the given quadratic equation. So, we get

$$\mathbf{X}^{\mathrm{T}}\mathbf{AX} + 8 = 0$$

or

$$(\mathbf{PX}')^{\mathrm{T}}\mathbf{A}(\mathbf{PX}') + 8 = 0$$

or

$$\mathbf{X}'^{\mathrm{T}}\mathbf{DX}' + 8 = 0 \qquad \text{where } \mathbf{D} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$$

Hence

$$[x' \ y'] \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + 8 = 0$$

or

$$-2x'^2 + 3y'^2 + 8 = 0$$

$\therefore$

$$\frac{x'^2}{4} - \frac{y'^2}{8/3} = 1$$

which is the equation of the hyperbola.

EXAMPLE 5.59 By using the rotation and translation of co-ordinate axes, put the quadratic

$$9x^2 - 4xy + 6y^2 - 10x - 20y - 5 = 0$$

in standard position. Also, give the name of the quadratic.

Solution: The matrix representation of the given quadratic equation is

$$9x^2 - 4xy + 6y^2 - 10x - 10y - 5 = [x \ y] \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [-10 \ -20] \begin{bmatrix} x \\ y \end{bmatrix} - 5 = 0 \qquad \text{(i)}$$

Take $\mathbf{A} = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$ and $\mathbf{k} = [-10 \quad -20]$

To construct a matrix **P** that is orthogonally diagonalized, we will try to find the orthonormal bases of each eigenspace of **A**. The characteristic equation of **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 9 & 2 \\ 2 & \lambda - 6 \end{bmatrix}\right) = 0$$

or $$\lambda^2 - 15\lambda + 50 = 0$$

or $$(\lambda - 10)(\lambda - 5) = 0$$

$\therefore$ $$\lambda = 5 \text{ or } \lambda = 10$$

Thus, $\lambda_1 = 5$ and $\lambda_2 = 10$ are the eigenvalues of **A**. It is easy to find that $\mathbf{w}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$ are the orthogonal bases for the eigenspaces corresponding to eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 10$ respectively.

So we can construct the matrix **P** such that det **P** = 1

Hence

$$\mathbf{P} = [\mathbf{w}_1 \quad \mathbf{w}_2] = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

Rotate the co-ordinate axes of the *xy*-co-ordinate system by substituting

$$\mathbf{X} = \mathbf{PX}'$$

i.e. $$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

in the given quadratic equation. So, we get

$$\mathbf{X}^{\mathrm{T}}\mathbf{AX} + k\mathbf{X} - 5 = 0$$

or $$(\mathbf{PX}')^{\mathrm{T}}\mathbf{A}(\mathbf{PX}') + k(\mathbf{PX}') - 5 = 0$$

or $$\mathbf{X}'^{\mathrm{T}}\mathbf{DX}' + k(\mathbf{PX}') - 5 = 0 \qquad \text{where } \mathbf{D} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

$$\therefore \quad [x' \ \ y']\begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} + [-10 \ \ -20]\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} - 5 = 0$$

or
$$5x'^2 + 10y'^2 - \frac{50}{\sqrt{5}}x' - 5 = 0$$

or
$$5[x' - \sqrt{5}]^2 + 10y'^2 - 30 = 0 \tag{ii}$$

To put the conic in standard position, translate the co-ordinates by substituting

$$x'' = x' - \sqrt{5}, \ \ y'' = y'$$

in Eq. (ii). So, we get

$$x''^2 + 2y''^2 = 6$$

which is the equation of ellipse.

Quadratic Surface in R^3

We consider here the quadratic equation of the form.

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + j = 0 \tag{5.2}$$

where $a, b, c, d, e, f, g, h, i$ and j are real numbers and a, b, c, d, e and f are not all zero. In the quadratic Eq. (5.2), the term

$$ax^2 + 2bxy + cy^2 + 2dxy + 2exz + 2fyz$$

is called the *associated quadratic form*.

Theorem 5.31 [Principal Axes Theorem in R^3]

If the equation of a quadratic surface in the xyz-co-ordinate system is

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + j = 0$$

and the associated quadratic form is

$$\mathbf{X}^{\mathrm{T}}\mathbf{AX} = ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz$$

then the co-ordinate axes of the xyz-co-ordinate system can be rotated on to the co-ordinate axes of the $x'y'z'$-co-ordinate system such that the equation of the quadratic surface is

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + g'x' + h'y' + i'z' + j = 0$$

where λ_1, λ_2 and λ_3 are the eigenvalues of $\mathbf{A}$. Moreover, the above rotation can be done by the change of variable method where $\mathbf{P}$ orthogonally diagonalizes $\mathbf{A}$ and det $\mathbf{P} = 1$.

EXAMPLE 5.60 By using the rotation and translation of co-ordinate axes, put the quadratic

$$2xy + 2xz + 2yz - 6x - 6y - 4z + 9 = 0$$

in standard position. Also, give the name of the quadratic.

Solution: The matrix form of the given quadratic is

$$2xy + 2xz + 2yz - 6x - 6y - 4z + 9 = 0$$

$$[x \;\; y \;\; z]\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} + [-6 \;\; -6 \;\; -4]\begin{bmatrix} x \\ y \\ z \end{bmatrix} + 9 = 0 \qquad \text{(i)}$$

Take $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ and $\mathbf{k} = [-6 \;\; -6 \;\; -4]$

To construct a matrix **P** that orthogonally diagonalizes a matrix **A**, we will find the orthonormal bases of each eigenspace of **A**.

The characteristic equation of **A** is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{bmatrix}\right) = 0$$

or $$\lambda^3 - 3\lambda - 2 = 0$$

or $$(\lambda + 1)^2 (\lambda - 2) = 0$$

$\therefore$ $$\lambda = -1, -1 \text{ or } \lambda = 2$$

Thus $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 2$ are the eigenvalues of **A**. It is easy to find that $\mathbf{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$,

$\mathbf{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$ and $\mathbf{w}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$ are the orthogonal bases for the eigenspaces corresponding to eigenvalues $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 2$ respectively.

Construct the matrix **P** such that det **P** = 1

So we can construct the matrix **P** such that det **P** = 1

$$\mathbf{P} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; \mathbf{w}_3] = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

First we rotate the co-ordinate axes of the *xy*-co-ordinate system by substituting

$$\mathbf{X} = \mathbf{PX}'$$

i.e.
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{P} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

in the given quadratic equation. So we get

$$\mathbf{X}^{\mathrm{T}}\mathbf{AX} + \boldsymbol{k}\mathbf{X} + 9 = 0$$

or
$$(\mathbf{PX}')^{\mathrm{T}}\mathbf{A}(\mathbf{PX}') + \mathbf{k}(\mathbf{PX}') + 9 = 0$$

or
$$\mathbf{X}'^{\mathrm{T}}\mathbf{DX} + \mathbf{k}(\mathbf{PX})' + 9 = 0 \qquad \text{where } \mathbf{D} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Hence

$$[x' \;\; y' \;\; z'] \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + [-6 \;\; -6 \;\; -4]\,\mathbf{P} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + 9 = 0$$

or
$$-x'^2 - y'^2 + 2z'^2 + \frac{4}{\sqrt{6}}y' - \frac{16}{\sqrt{3}}z' + 9 = 0$$

or
$$x'^2 + y'^2 - 2z'^2 - \frac{4}{\sqrt{6}}y' + \frac{16}{\sqrt{3}}z' - 9 = 0$$

or
$$x'^2 + \left(y' - \frac{2}{\sqrt{6}}\right)^2 - 2\left(z' - \frac{4}{\sqrt{3}}\right)^2 + 1 = 0 \qquad \text{(ii)}$$

To put the conic in standard position, translate the co-ordinates by substituting

$$x'' = x', \qquad y'' = y' - \frac{2}{\sqrt{6}}, \qquad z'' = z' - \frac{4}{\sqrt{3}}$$

in Eq. (ii). So, we get

$$x''^2 + y''^2 - 2z''^2 = -1$$

which is a hyperboloid of two sheets.

SUMMARY

Eigenvalues and Eigenvectors A vector $x \in R^n$ is called an *eigenvector* of a square matrix $\mathbf{A} = [a_{jk}]$ of order n if

$$\mathbf{Ax} = \lambda \mathbf{x} \qquad \text{for some scalar } \lambda,$$

and λ is called *eigenvalue* of **A**.

Characteristic Equation Let **A** be a square matrix and λ be a scalar. Then the scalar equation

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

where **I** is identity matrix of same order as that of matrix **A**, is called the *characteristic equation* of **A**.

Theorem [Eigenvalues] The eigenvalues of a square matrix **A** are the roots of the characteristic equation of **A**.

Remarks:

(i) A square matrix of order n has at least one eigenvalue and at most n numerically different eigenvalues.
(ii) If **x** is an eigenvector of a matrix **A** corresponding to an eigenvalue λ, then $k\mathbf{x}$ $(k \neq 0)$ is also an eigenvector corresponding to the eigenvalue λ.

Theorem [Eigenvalues and Eigenvectors] Let **A** be a square matrix of order n and λ be a real number. Then the following statements are equivalent.

(i) λ is an eigenvalue of **A**
(ii) The system of equation $(\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ has non-trivial solutions.
(iii) There is a nonzero vector **x** in R^n such that $\mathbf{Ax} = \lambda \mathbf{x}$.
(iv) λ is a solution of the characteristic equation $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$.

Theorem [Eigenvalues of a lower Triangular, Upper Triangular or Diagonal Matrix] If a square matrix **A** is a lower triangular (or upper triangular or diagonal) matrix, then the eigenvalues of **A** are the entries on the main diagonal of **A**.

Eigenspaces If **A** is a square matrix of order n, then the solution space of characteristic equation

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

is called the eigenspace of **A** corresponding to the eigenvalue λ.

Theorem If $\lambda_1, \lambda_2, \ldots, \lambda_n$ (may not be distinct) are the eigenvalues of a square matrix of order n, then:

(i) The determinant of **A** is $\det \mathbf{A} = \lambda_1 \lambda_2 \ldots \lambda_n$
(ii) The trace of **A** is $\operatorname{tr} \mathbf{A} = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Corollary A square matrix **A** is invertible if and only if $\lambda = 0$ is not an eigenvalue of **A**.

Corollary If the characteristic equation of **A** is $\lambda^n + C_1 \lambda^{n-1} + \cdots + C_n$, then $\det \mathbf{A} = (-1)^n C_n$.

Remark From the above two corollaries, a square matrix **A** is invertible if and only if $C_n \neq 0$.

Theorem If λ is an eigenvalue of a square matrix **A** corresponding to an eigenvector **x**, then for a positive integer k, λ^k is an eigenvalue of $\mathbf{A}^k$ corresponding to an eigenvector **x**.

Result If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of a matrix **A**, then the trace of $\mathbf{A}^k$ is:

$$\mathbf{A}^k = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k.$$

Result

(i) If λ is an eigenvalue of a square matrix **A** corresponding to an eigenvector **x**, then for a scalar C, $\lambda - C$ is an eigenvalue of $\mathbf{A} - C\mathbf{I}$ corresponding to the eigenvector **x**.

(ii) Let **A** be an invertible matrix. If λ is an eigenvalue of **A** corresponding to an eigenvector **x**, then $\dfrac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$ corresponding to the eigenvector **x**.

Theorem [Cayley–Hamilton Theorem] Every square matrix **A** satisfies its characteristic equation. That is, if

$$a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = 0$$

is the characteristic equation of **A**, then

$$a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + \cdots + a_n\mathbf{A}^n = \mathbf{0}.$$

Theorem [Eigenvalues of Real Symmetric Matrices] The eigenvalues of a real symmetric matrix are real.

Theorem [Eigenvalues of Real Skew-Symmetric Matrices] The eigenvalues of a real skew-symmetric matrix are pure imaginary or zero.

Result The value of determinant of an orthogonal matrix is 1 and –1.

Theorem [Orthogonal Matrix] A real square matrix is orthogonal if and only if the rows of the matrix form an orthonormal set of vectors, or if and only if the columns of the matrix form an orthonormal set of vectors.

Theorem [Eigenvalues of Orthogonal Matrices] The eigenvalues of an orthogonal matrix are real or complex conjugate in pairs and have the absolute value 1.

Complex Matrices A matrix with complex entries is called *complex matrix*. Here we will discuss some special complex matrices: Hermitian, skew-Hermitian, unitary and normal matrices.

Conjugate of Complex Matrix If $\mathbf{A} = [a_{ij}]$ is a complex matrix, then the conjugate of **A** is $\overline{A} = [\overline{a}_{ij}]$ where $\overline{a}_{ij}$ is a complex conjugate of **A**.

Hermitian Matrix If $\mathbf{A} = [a_{ij}]$ is a complex matrix and

$$\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{A} \qquad \text{where } \overline{\mathbf{A}}^{\mathrm{T}} = \text{conjugate transpose of } \mathbf{A},$$

that is

$$a_{ij} = \overline{a}_{ji} \qquad \text{for all } i \text{ and } j,$$

then **A** is called the Hermitian matrix.

Skew-Hermitian Matrix If $\mathbf{A} = [a_{ij}]$ is a complex matrix and

$$\overline{\mathbf{A}}^{\mathrm{T}} = -\mathbf{A} \qquad \text{where } \overline{\mathbf{A}}^{\mathrm{T}} = \text{conjugate transpose of } \mathbf{A},$$

that is

$$a_{ij} = -\overline{a}_{ji} \qquad \text{for all } i \text{ and } j,$$

then **A** is called skew- Hermitian matrix.

Remarks:

(i) If **A** is a Hermitian matrix, then the entries on the main diagonal must satisfy $a_{ii} = \overline{a}_{ii}$, that is, they are real.

Theorem [Geometric Multiplicity] Let **A** be an $n \times n$ matrix. Then the geometric multiplicity is less than or equal to the algebraic multiplicity for every eigenvalue of **A**.

Theorem [Diagonalization of an $n \times n$ Matrix] Let **A** be an $n \times n$ matrix. Then **A** is diagonalizable if and only if the geometric multiplicity is equal to the algebraic multiplicity for every eigenvalue of **A**.

Theorem [Powers of a Diagonalizable Matrix] Let **A** be an $n \times n$ diagonalizable matrix and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, then

(i) For any positive integer k,

$$\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}.$$

(ii) If all the diagonal entries of **D** are nonzero, then **A** is invertible and

$$\mathbf{A}^{-1} = \mathbf{P}\mathbf{D}^{-1}\mathbf{P}^{-1}$$

Orthogonal Diagonalization A real square matrix **A** is called *orthogonally diagonalizable* if there exists an orthogonal matrix **P** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ ($= \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}$) is diagonal and **P** is then said to orthogonally diagonalize **A**.

Theorem [Orthogonally Diagonalized Matrix] An $n \times n$ real matrix **A** is orthogonally diagonalized if **A** is symmetric.

Unitary Diagonalization A complex square matrix **A** is called *unitary diagonalizable* if there exists a unitary matrix **P** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ ($=\mathbf{P}^*\mathbf{A}\mathbf{P}$) is diagonal and **P** is then said to unitarily diagonalize **A**.

Remark An $n \times n$ complex matrix **A** is *unitary* if and only if the row vectors (or column vectors) of **A** form an orthonormal set in C^n.

Theorem [Unitary Diagonalization] An $n \times n$ complex matrix **A** is unitarily diagonalized if **A** is normal.

Theorem [The Spectral Theorem for Real Symmetric Matrix] Every real symmetric matrix is orthogonally diagonalizable.

Quadratic Form The quadratic form in n variables $x_1, x_2, x_3, \ldots, x_n$ is

$$\begin{aligned}&a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ &+ \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2\end{aligned}$$

where all the coefficient a_{ij} $(1 \le i \le j \le n)$ are real scalars. Moreover, the terms of the form $a_{ij}x_ix_j$ $(i \neq j)$ in a quadratic form are called *cross product terms*.

Remark The quadratic form can be expressed as a matrix multiplication $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}$

$$\text{where } \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Classification of Quadratic Forms Let **A** be a symmetric matrix. The quadratic form is $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}$ called

(a) positive definite if $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} > 0$ for all $x \neq 0$
(b) semi-positive definite if $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} \geq 0$ for all x
(c) negative definite if $\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} < 0$ for all $x \neq 0$

(d) semi-negative definite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX} \le 0$ for all x
(e) indefinite if $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ has both positive and negative values

Theorem [Symmetric Matrix] For any $n \times n$ matrix $\mathbf{A}$ and any $1 \times n$ row matrix $\mathbf{X}$, we have $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \mathbf{X}^{\mathrm{T}}\mathbf{BX}$ where $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$ is the symmetric matrix.

Theorem [Characterization of Positive Definite Form] Let $\mathbf{A}$ be a symmetric matrix. Then the following statements are equivalent.

(a) The quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is positive definite.
(b) The matrix $\mathbf{A}$ is positive definite.
(c) All the eigenvalues of $\mathbf{A}$ are positive.

Principal Submatrix Let $\mathbf{A}$ be an $n \times n$ square matrix. Then the *principal submatrix* $\mathbf{A}_r$ of order r of $\mathbf{A}$ is the submatrix formed from the first r rows and r columns of $\mathbf{A}$ for $r = 1, 2, \ldots, n$.

Theorem [Alternative Method of Characterization of Positive Definite Form] Let $\mathbf{A}$ be a symmetric matrix. Then the following statements are equivalent.

(a) The quadratic form $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ is positive definite.
(b) The matrix $\mathbf{A}$ is positive definite.
(c) The determinant of every principal submatrix is positive.

Remark The matrix $\mathbf{A}$ will be

(i) positive definite if $\det(\mathbf{A}_i) > 0$ for all i.
(ii) positive semi-definite if $\det(\mathbf{A}_i) \ge 0$ for all i. ($\det(\mathbf{A}_i) = 0$ for at least one i)
(iii) negative definite if $\det(\mathbf{A}_1)$, $\det(\mathbf{A}_3)$, $\det(\mathbf{A}_5)\ldots$ are negative and $\det(\mathbf{A}_2)$, $\det(\mathbf{A}_4)$, $\det(\mathbf{A}_6)\ldots$ are positive.
(iv) negative semi-definite if $\det(\mathbf{A}_i) = 0$ for at least one i in case (iii).
(v) indefinite in all other cases.

Theorem [Maximum and Minimum Values of Quadratic Form] Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. If the eigenvalues of $\mathbf{A}$ are $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ and $\|x\| = 1$, then

(i) $\lambda_1 \ge \mathbf{X}^{\mathrm{T}}\mathbf{AX} \ge \lambda_n$
(ii) $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \lambda_1$ if $\mathbf{X}$ is an eigenvector corresponding to λ_1 and $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = \lambda_n$ if $\mathbf{X}$ is an eigenvector corresponding to λ_n.

Diagonalization of a Quadratic Form The diagonalization of a quadratic form is a process of removing the cross product terms form a quadratic form. In other words, it is the process to reduce the quadratic form into a sum of squares.

Theorem [Diagonalization of Quadratic Forms] Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. If there is a change of variable by substituting $\mathbf{X} = \mathbf{PY}$, then it transforms the quadratic $\mathbf{X}^{\mathrm{T}}\mathbf{AX}$ into a quadratic form $\mathbf{Y}^{\mathrm{T}}\mathbf{AY}$ with no cross product terms.

Theorem [Principal Axes Theorem in R^2] If the equation of the conic in the xy-co-ordinate system is

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

and the associated quadratic form is $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = ax^2 + 2bxy + cy^2$, then the co-ordinate axes of the

xy-co-ordinate system can be rotated on to the co-ordinate axes of $x'y'$-co-ordinate system such that the equation of the conic in the new $x'y'$-co-ordinate system is

$$\lambda_1 x'^2 + 2bx'y' + \lambda_2 y'^2 + d'x' + e'y' + f = 0$$

where λ_1 and λ_2 are eigenvalues of **A**. Moreover, the above rotation can be done by the change of variable method where **P** orthogonally diagonalizes **A** and det **P** = 1.

Theorem [Principal Axes Theorem in R^3] If the equation of a quadratic surface in the xyz-co-ordinate system is

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + gx + hy + iz + j = 0$$

and the associated quadratic form is $\mathbf{X}^{\mathrm{T}}\mathbf{AX} = ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz$, then the co-ordinate axes of the xyz-co-ordinate system can be rotated on to the co-ordinate axes of $x'y'z'$-co-ordinate system such that the equation of the quadratic surface is

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + g'x' + h'y' + i'z' + j = 0$$

where λ_1, λ_2 and λ_3 are the eigenvalues of **A**. Moreover, the above rotation can be done by the change of variable method where **P** orthogonally diagonalizes **A** and det **P** = 1.

Answers to Exercise Sets

CHAPTER 1

Answers to Exercise Set 1

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} = \begin{bmatrix} 5 & 2 & 7 \\ 4 & 4 & -2 \end{bmatrix}$; $(-4)\ \mathbf{B} = \begin{bmatrix} -4 & 0 & -8 \\ -12 & -4 & -16 \end{bmatrix}$; $6\mathbf{A} = \begin{bmatrix} 24 & 12 & 30 \\ 6 & 18 & -36 \end{bmatrix}$;

$5\mathbf{A} - 3\mathbf{A} = \begin{bmatrix} 17 & 10 & 19 \\ -4 & 12 & -37 \end{bmatrix}$; $2\mathbf{B} - 3\mathbf{B} = \begin{bmatrix} -1 & 0 & -2 \\ -3 & -1 & -4 \end{bmatrix}$.

4. $\mathbf{A}^2\mathbf{B} + \mathbf{B}^2\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

5. $x = -3, y = -2, z = 4, a = 3$.

6. $\mathbf{AB} = \begin{bmatrix} 3 & -2 \\ 5 & -5 \\ 7 & -8 \end{bmatrix}$; $\operatorname{tr} \mathbf{A} = 6$

7. $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 2 & -1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} -1 & -2 \\ 4 & -1 \end{bmatrix}$

8. $\mathbf{AB} = \begin{bmatrix} -2 & -2 \\ 0 & -6 \end{bmatrix}$

11. (i), (ii), (v) are in row-echelon form while (iii) and (iv) are not in row-echelon form.

12. (i), (ii), (iii) are in row-echelon form while (iv) and (v) are not in row-echelon form.

13. (i) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 8 & -11 \\ 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -\frac{3}{2} \\ 0 & 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$

14. (i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 22 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Answers to Exercise Set 2

3. (i), (iii) are symmetric matrices while (ii) is not a symmetric matrix.

4. (i), (ii) are skew-symmetric matrices while (iii) is not a skew-symmetric matrix.

6. $\mathbf{A} = \mathbf{B} + \mathbf{C}$ where $\mathbf{B} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & -2 \\ 3 & -2 & 2 \end{bmatrix}$; $\mathbf{C} = \begin{bmatrix} 0 & -1 & -3 \\ 1 & 0 & 6 \\ 3 & -6 & 0 \end{bmatrix}$.

Answers to Exercise Set 3

1. (i) , (ii), (v) are elementary matrices while (iii) and (iv) are not elementary matrices.

2. (i) $\begin{bmatrix} -7 & 4 \\ 2 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} -\frac{5}{39} & \frac{2}{13} \\ \frac{4}{13} & \frac{1}{13} \end{bmatrix}$

(iii) An inverse of the given matrix does exist.

(iv) $\begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$.

5. $\mathbf{B}^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$

Answers to Exercise Set 4

1. (i) 5 (ii) $-2a$ (iii) 165

(iv) -546 (v) -12

2. $-\frac{1}{2}\begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix}$

3. (i) $\frac{1}{64}\begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix}$ (ii) $\left(-\frac{1}{3}\right)\begin{bmatrix} 6 & 6 & -15 \\ 1 & 0 & -1 \\ -5 & -3 & 8 \end{bmatrix}$ (iii) $\frac{1}{20}\begin{bmatrix} 2 & 6 & 4 \\ 21 & -7 & -8 \\ -18 & 6 & 4 \end{bmatrix}$

(iv) $\frac{1}{152}\begin{bmatrix} 29 & 11 & -19 \\ -21 & 13 & 19 \\ 27 & 5 & 19 \end{bmatrix}$

7. $n-1$

Answers to Exercise Set 5

1. (i) 2 (ii) 3 (iii) 2
(iv) 4 (v) 2 (vi) 4
(vii) 4 (viii) 2

2. (b) 1 **3.** (a) 0

4. (c) 3 **5.** (b) 0

Answers to Exercise Set 6

1. (i) $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$

2. (i) not consistent (ii) consistent (iii) consistent
(iv) consistent

3. $\lambda \neq \frac{7}{10}, \mu = \frac{7}{10}$

4. (i) $a = 8, b \neq 18$ (ii) $a \neq 8$, for any b

5. (i) $\lambda = 5, \mu \neq 9$ (ii) $\lambda \neq 5$, μ arbitrary (iii) $\lambda = 5,\ \mu = 9$.

6. (i) $a = 8, b \neq 15$ (ii) $a \neq 8$, b arbitrary (iii) $a = 8, b = 15$.

7. (i) $k = 1$ (ii) $k = 1$

Answers to Exercise Set 7

1. $x = 1, y = -1,\ z = 1$ **2.** $x = -\frac{3}{7}, y = \frac{8}{7}, z = -\frac{2}{7}$

3. $x = 2, y = 1, z = -1$ **4.** $x = 2, y = 1, z = 0$

5. $x = 3, y = 5, z = 6$ **6.** $x = 0, y = -1, z = 1$

7. $x_1 = \frac{396}{175},\ x_2 = \frac{24}{25},\ x_3 = \frac{72}{175}$ **8.** $x_1 = 2, x_2 = 1, x_3 = -1, x_4 = 3$

Answers to Exercise Set 8

1. (i) $x = 2, y = -3$ (ii) $x = \frac{1}{3}, y = 1, z = -\frac{1}{3}$ (iii) Inconsistent system

(iv) $x = 1, y = 3t - 2, z = t$ (v) $x = 1, y = -1, z = -1$

(vi) $x_1 = 10,\ x_2 = 20,\ x_3 = 40,\ x_4 = 50$

2. (i) $\mu = 3$ or 0 (ii) $\mu = \frac{1}{4}$

3. $x = 1.2, y = 2.2, z = 3.2$ **4.** $x = k, y = 2k, z = k$

Answers to Exercise Set 9

(i) $x = 2, y = -3$ (ii) $x = 3, y = 4, z = 8$ (iii) $x = -3, y = 4,\ z = 1$

(iv) $x = 1.5, y = 1, z = 2.5$ (v) $x = 2, y = 1, z = -4$ (vi) $x = z = k, y = 2k$

(vii) $x = 1 + t, y = 1 + 14t, z = 5t, w = 0$

(viii) $x_1 = 2, x_2 = 1, x_3 = -1, x_4 = 3$

Answers to Exercise Set 10

(i) $x = -1, y = 4$ (ii) $x = 1,\ y = 3,\ z = 5$ (iii) $x = 2, y = -1, z = \frac{1}{2}$

(iv) $x = 2, y = 1, z = -4$ (v) $x = -3, y = -7$ (vi) $x = 1 - 3t, y = -1 + 4t, z = t$

(vii) Inconsistent system (viii) $x_1 = 2, x_2 = 1, x_3 = -1, x_4 = 3$ (ix) Inconsistent system

(x) $x = 0, y = 0, z = 0$ (xi) $x = 1, y = 2, z = 1$

CHAPTER 2

Answers to Exercise Set 1

1. (i) $(9, 5)$ (ii) $\sqrt{106}$ (iii) $\left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$

(iv) 1 (v) $(-5, -7)$.

2. (i) $(5, -1, 3)$ (ii) $(2, -6, 4)$ (iii) $(-4, -2, -1)$

(iv) $(6, -1, -4)$

3. (i) $\mathbf{v} - \mathbf{w} = (-2, 1, -4)$ (ii) $6\mathbf{u} + 2\mathbf{v} = (-10, 6, -4)$ (iii) $-\mathbf{v} + \mathbf{u} = (-7, 1, 10)$

(iv) $5(\mathbf{v} - 4\mathbf{u}) = (80, -20, -80)$ (v) $-3(\mathbf{v} - 8\mathbf{w}) = (132, -24, -72)$

4. (i) $\sqrt{83}$ (ii) $\sqrt{17} + \sqrt{26}$ (iii) $4\sqrt{17}$

5. $k = \pm \frac{2\sqrt{2}}{\sqrt{15}}$ **6.** $\theta = 60$

7. (i) orthogonal (ii) obtuse (iii) acute

(iv) acute.

8. $c = (-3)$ **9.** $k = -4$

10. $(-4, 7, 21), (5, 11, 21), \sqrt{13}$

14. (i), (ii), (iii), (iv) and (vi) have satisfied the Pythagoras theorem while (v) has not satisfied the Pythagoras theorem.

Answers to Exercise Set 2

1. (i) is a vector space. (ii) **B** is a vector space while **A** is not a vector space. (iii) is not a vector space.

3. It is not a vector space.

7. (i) is a subspace while (ii) and (iii) are not subspaces.

8. It is a subspace.

9. (ii) is a subspace while (i) (iii) and (iv) are not subspaces.

10. (i) and (ii) can be expressed as linear combination of the given vectors while (iii) can not be.

14. It cannot be spanned R^3.

15. (i) and (ii) spanned R^3 while (iii) and (iv) cannot be spanned R^3.

Answers to Exercise Set 3

1. (i), (iii), (v) and (vi) are linearly independent sets while (ii), (iv) and (vii) are linearly dependent sets.

2. (iii) and (iv) are linearly independent sets while (i) and (ii) are linearly dependent sets. (ii) and (iii) are linearly independent sets while (i), (iv) and (v) are linearly dependent sets.

Answers to Exercise Set 4

1. (i), (iv) and (v) are bases but (ii), (iii) are not bases for their indicated spaces.

2. (i) $(\mathbf{v})_S = \left(\frac{23}{10}, \frac{1}{2}, \frac{3}{5}\right)$ (ii) $(\mathbf{p})_S = (-2, 3, 4)$ (iii) $(M)_S = (-5, -1, 0, -2)$

3. (i) $\dim(S) = 3$ (ii) $\dim(S) = 3$.

4. Extended basis = $\{(1, -1, 0), (3, 1, -2), (1, 0, 0)\}$ or $\{(1, -1, 0), (3, 1, -2), (0, 1, 0)\}$.

Answers to Exercise Set 5

1. (i) Row vectors: $\mathbf{r}_1 = [1,\ 3]$, $\mathbf{r}_2 = [-2,\ 4]$;

Column vectors: $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

(ii) Row vectors: $\mathbf{r}_1 = [0\ \ 2\ \ 3]$, $\mathbf{r}_2 = [-5\ \ 6\ \ 1]$; $\mathbf{r}_3 = [3\ \ 2\ \ -4]$

Column vectors: $\mathbf{c}_1 = \begin{bmatrix} 0 \\ -5 \\ 3 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{c}_3 = \begin{bmatrix} 3 \\ 1 \\ -4 \end{bmatrix}$.

(iii) Row vectors: $\mathbf{r}_1 = [-1\ \ 2\ \ 6\ \ 0]$, $\mathbf{r}_2 = [3\ \ -7\ \ -8\ \ 9]$;

Column vectors: $\mathbf{c}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 2 \\ -7 \end{bmatrix}$, $\mathbf{c}_3 = \begin{bmatrix} 6 \\ -8 \end{bmatrix}$, $\mathbf{c}_3 = \begin{bmatrix} 0 \\ 9 \end{bmatrix}$

2. (i) $R(\mathbf{A}) = \text{span}\,\{\mathbf{r}_1,\ \mathbf{r}_2\} = \{\alpha_1[-2\ \ 4] + \alpha_2[1\ \ 3] \mid \alpha_1, \alpha_2 \in R\}$,

$$C(\mathbf{A}) = \text{span}\,\{\mathbf{c}_1,\ \mathbf{c}_2\} = \left\{\beta_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \beta_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} \middle|\ \beta_1, \beta_2 \in R\right\},$$

(ii) $R(\mathbf{A}) = \text{span}\,\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$; $C(\mathbf{A}) = \text{span}\,\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$

(iii) $R(\mathbf{A}) = \text{span}\,\{\mathbf{r}_1, \mathbf{r}_2\}$; $C(\mathbf{A}) = \text{span}\,\{\mathbf{c}_1, \mathbf{c}_2\}$.

3. Basis for row space $\left\{(1, -2, 0, 0, -11), \left(0, 1, -\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, -\frac{3}{8}\right), (0, 0, 1, 2, 0, -5)\right\}$

Basis for column space $= \left\{\begin{bmatrix} 2 \\ -1 \\ 10 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix},\right\}$

4. Basis for row space = $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_4\}$ where $\mathbf{r}_1 = [1\ \ -2\ \ 0\ \ 3]$, $\mathbf{r}_2 = [2\ \ -5, -3\ \ 6]$, $\mathbf{r}_4 = [2\ \ -1\ \ 4\ \ -7]$.

5. Basis for space spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ = $\{(1, -2, 1, -5, -6), (0, 1, -2, 2, 4), (0, 0, 0, 1, -2)\}$.

6. Basis for the space spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$= $\{v_1, v_2, v_4\}$.

7. Null space = $\{s(-1, 1, 0, 0, 0) + t(-1, 0, -1, 0, 1) \mid s, t \in R\}$.

8. Basis for null space $= \left\{\left(-1, -\frac{1}{2}, -2, 1, 0, 0\right), \left(0, -\frac{1}{2}, -2, 1, 0, 0\right), (1, 1, 5, 0, 0, 1)\right\}$

9. (i) $\mathbf{b} = \begin{bmatrix} 4 \\ 13 \\ -1 \end{bmatrix} = (-1)\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$

(ii) $\mathbf{b}$ is not a vector of column space of $\mathbf{A}$.

10. Solution $= \begin{bmatrix} -2 \\ 7 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$; General solution of the homogeneous system $= t\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$;

Particular solution of non-homogeneous system $= \begin{bmatrix} -2 \\ 7 \\ 0 \end{bmatrix}$.

Answer to Exercise Set 6

1. rank $\mathbf{A}$ = 3 and nullity $\mathbf{A}$ = 2.

2. rank $\mathbf{A}$ = 3 and nullity $\mathbf{A}$ = 3.

4. Maximum values of ranks of $\mathbf{A}$ and $\mathbf{B}$ are both 3.

CHAPTER 3

Answers to Exercise Set 1

1. (i) not linear (ii) linear (iii) not linear

2. (i) linear (ii) linear (iii) linear

3. (i) linear (ii) linear (iii) linear

4. (i) $y' = y$, $z' = z\cos\theta, - x\sin\theta$, $x' = z\sin\theta + x\cos\theta$ and it is linear.

(ii) $x' = x$, $y' = y\cos\theta - z\sin\theta$, $z' = z\cos\theta + y\sin\theta$ and it is linear.

5. (i) $T(x, y, z) = (x + z, 0, x + y)$, $T(1, 0, 1) = (2, 0, 1)$

(ii) $T(x, y) = (x, y, x + y)$, $T(2, 3) = (2, 3, 5)$

(iii) $T(a_0 + a_1x + a_2x^2) = a_0 + 2a_1 + 3a_2$, $T(5 - 3x + 2x^2) = 5$

(iv) $T(x, y) = (2x - y, x + y)$, $T(3, 3) = (3, 6)$

(v) $T(x, y, z) = (x + y + z, 2y, x - y - z)$, $T(1, -2, 1) = (0, -4, 2)$.

(vi) $T\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right\} = 3a - 4b + c - d \cdot T\left\{\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}\right\} = 6.$

Answers to Exercise Set 2

1. $\begin{bmatrix} 3 & 2 \\ 1 & -4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & -4 \\ 1 & 2 \\ 6 & -1 \\ 0 & 10 \end{bmatrix}$

3. (i) $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (iii) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4. (i) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

5. (i) $\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 1 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$ (ii) $\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 & -2 \\ 4 & -7 \end{bmatrix}$

9. (i) 3 (ii) 7
(iii) 6 (iv) –32.

Answers to Exercise Set 2

1. (i) $\langle \mathbf{u}, \mathbf{v} \rangle = 4;\quad \|\mathbf{u}\| = \sqrt{21};\quad d(\mathbf{u}, \mathbf{v}) = \sqrt{26}$

(ii) $\langle \mathbf{u}, \mathbf{v} \rangle = 0;\quad \|\mathbf{u}\| = \sqrt{\dfrac{86}{5}};\quad d(\mathbf{u}, \mathbf{v}) = \sqrt{\dfrac{296}{5}}$

(iii) $\langle \mathbf{u}, \mathbf{v} \rangle = \dfrac{1}{4};\quad \|\mathbf{u}\| = \dfrac{1}{\sqrt{3}};\quad d(\mathbf{u}, \mathbf{v}) = \dfrac{1}{\sqrt{30}}$

4. $\theta = \cos^{-1}\left(\dfrac{6}{\sqrt{3}\sqrt{14}}\right)$ **5.** $\theta = \dfrac{\pi}{2}$

6. Ellipsoid: $\dfrac{x^2}{1/4} + \dfrac{y^2}{1/9} + \dfrac{z^2}{1/16} = 1$

Answers to Exercise Set 3

1. (ii) and (iii) are orthogonal vectors but (i) and (iv) are not orthogonal vectors

3. (i) $k = 8$ (ii) $k = 3, -4$

4. A set of orthogonal vectors: $\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} t \;\middle|\; t \in R \right\}$.

5. A set of orthogonal vectors: $\left\{ \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} y + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} z \;\middle|\; y, z \in R \right\}$. Yes.

6. Orthogonal **7.** Orthogonal

11. $\left\{ [0, 1, 0], \left[\dfrac{1}{2}, 0, 1\right] \right\}$ **12.** $x + 3y = 0$.

13. $x - z = 0$.

Answers to Exercise Set 4

4. Yes **5.** Yes. $\left\{ \dfrac{1}{\sqrt{2\pi}}, \dfrac{\cos x}{\sqrt{\pi}}, \dfrac{\sin 2x}{\sqrt{\pi}} \right\}$

6. $(\mathbf{u})_S = [-2\sqrt{2}, 5\sqrt{2}]$

7. (i) $(0, 3)$, component of $\mathbf{u}$ orthogonal to $\mathbf{e}_1 = (2, 0)$

(ii) $\left(\frac{8}{5}, \frac{16}{5}\right)$, component of $\mathbf{u}$ orthogonal to $\mathbf{w} = \left(\frac{2}{5}, -\frac{1}{5}\right)$

8. $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$ where $\mathbf{w}_1 = \left(\frac{4}{25}, 1, -\frac{3}{25}\right) \in W$; $\mathbf{w}_2 = \left(\frac{21}{25}, 0, \frac{28}{25}\right) \in W^{\perp}$

9. $\left\{[2, -1, 0], \left[\frac{1}{5}, \frac{2}{5}, -1\right], \left[\frac{8}{3}, \frac{16}{3}, \frac{8}{3}\right]\right\}$

10. $\left\{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

11. $\left\{\frac{3}{\sqrt{50}} + \frac{4}{\sqrt{50}}x + \frac{5}{\sqrt{50}}x^2, \frac{3}{\sqrt{50}} + \frac{4}{\sqrt{50}}x - \frac{5}{\sqrt{50}}x^2, \frac{4}{5} - \frac{3}{5}x + 0x^2\right\}$

12. $\mathbf{Q} = \begin{bmatrix} \frac{2}{\sqrt{15}} & \frac{7}{\sqrt{4485}} & -\frac{118}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{64}{\sqrt{4485}} & \frac{11}{\sqrt{20930}} \\ -\frac{3}{\sqrt{15}} & -\frac{18}{\sqrt{4485}} & -\frac{81}{\sqrt{20930}} \\ \frac{1}{\sqrt{15}} & -\frac{4}{\sqrt{4485}} & -\frac{18}{\sqrt{20930}} \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} \sqrt{15} & -\frac{11}{\sqrt{15}} & \frac{17}{\sqrt{15}} \\ 0 & \frac{229}{\sqrt{4485}} & -\frac{53}{\sqrt{4485}} \\ 0 & 0 & \frac{210}{\sqrt{20930}} \end{bmatrix}$

13. $\mathbf{Q} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} & \frac{2}{\sqrt{6}} \\ 0 & -\frac{5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} \sqrt{5} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & \frac{6}{\sqrt{30}} & \frac{22}{\sqrt{30}} \\ 0 & 0 & \frac{16}{\sqrt{6}} \end{bmatrix}$

Answers to Exercise Set 5

1. $y = -12.1x + 29.4$

2. $\overline{\mathbf{x}} = \begin{bmatrix} =\frac{18}{7} \\ -\frac{151}{210} \\ \frac{107}{210} \end{bmatrix}$

3. $\text{proj}_{\text{w}}\, \mathbf{u} = (-2, 3, 4, 0)$

12. If the field of scalars is the set of real numbers R, then real eigenvalues exist only when $\sin\theta = 0$, in which case there are two equal eigenvalues $\lambda_1 = \lambda_2 = \cos\theta$, where $\cos\theta = 1$ or -1. In this case every nonzero vector is an eigenvalue for $\lambda_1 = 1$ and $\lambda_2 = -1$. If the field of scalars is the set of complex numbers C, then the real eigenvalues are $\lambda_1 = \cos\theta + i\sin\theta = 0$ and $\lambda_2 = \cos\theta - i\sin\theta = 0$. If $\sin\theta = 0$ these are real and equal. If $\sin\theta \neq 0$, then they are distinct complex conjugates; the eigenvectors belonging to λ_1 are $a[i, 1]$, $a \neq 0$; those belonging to λ_2 are $b[1, i]$, $b \neq 0$.

13. $\operatorname{tr}\mathbf{A}^2 = 1 + 4 + 16 = 21$. $\det \mathbf{A}^{-1} = 1 \cdot \dfrac{1}{2} \cdot \dfrac{1}{4} = \dfrac{1}{8}$.

Answers to Exercise Set 2

6. (i) $\mathbf{A}^{-1} = \dfrac{1}{2}\begin{bmatrix} 1-i & 1+i \\ 1+i & 1-i \end{bmatrix}$ (ii) $\mathbf{A}^{-1} = \dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ -i & -1 \end{bmatrix}$ (iii) $\mathbf{A}^{-1} = \begin{bmatrix} \dfrac{1-i}{2} & -\dfrac{i}{\sqrt{3}} & \dfrac{3-i}{2\sqrt{15}} \\ -\dfrac{1}{2} & \dfrac{1}{\sqrt{3}} & \dfrac{4-3i}{2\sqrt{15}} \\ \dfrac{1}{2} & \dfrac{i}{\sqrt{3}} & -\dfrac{5i}{2\sqrt{15}} \end{bmatrix}$.

9. Hermitian matrix: $\begin{bmatrix} 1 & 1-i & 2-3i \\ 1+i & 2 & i \\ 2+3i & -i & 7 \end{bmatrix}$; skew-Hermitian matrix: $\begin{bmatrix} i & 1+i & 3-2i \\ -1+i & i & 4+i \\ -3-2i & -4+i & 0 \end{bmatrix}$

Answers to Exercise Set 3

1. Yes, $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

2. $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -5 & 18 \\ -3 & 10 \end{bmatrix}$

3. (i) A matrix $\mathbf{A}_1$ cannot be diagonalized matrix because it has only one line of eigenvectors.

(ii) A matrix $\mathbf{A}_2$ cannot be diagonalized matrix because it has only two linearly independent eigenvectors.

4. (i) $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ (ii) $\begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$ (iii) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

(iv) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ (v) $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & \dfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ (vi) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

5. (i) $\mathbf{P} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

(ii) $\mathbf{P} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. $\mathbf{A}^4 = \begin{bmatrix} -49 & -50 & -40 \\ 65 & 66 & 40 \\ 130 & 130 & 81 \end{bmatrix}$.

7. (a) True: det $\mathbf{A} = 2 \neq 0$. (b) False: $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ (c) False: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

8. (a) False: Do not know λ's. (b) True (c) True
(d) False: need eigenvectors of S.

9. $\mathbf{A}$ is a diagonal matrix since $\mathbf{P} = \mathbf{I}$, $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} = \mathbf{I}^{-1}\mathbf{A}\,\mathbf{I} = \mathbf{A}$.

Answers to Exercise Set 4

1. Yes. $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

2. (i) $\mathbf{P} = \dfrac{1}{5}\begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathrm{AP} = \begin{bmatrix} 0 & 0 \\ 0 & 25 \end{bmatrix}$

(ii) $\mathbf{P} = \begin{bmatrix} \dfrac{1}{\sqrt{10}} & \dfrac{3}{\sqrt{14}} & \dfrac{3}{\sqrt{35}} \\ 0 & \dfrac{2}{\sqrt{14}} & -\dfrac{5}{\sqrt{35}} \\ \dfrac{3}{\sqrt{10}} & -\dfrac{1}{\sqrt{14}} & -\dfrac{1}{\sqrt{35}} \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

(iii) $\mathbf{P} = \dfrac{1}{\sqrt{50}}\begin{bmatrix} 0 & 5 & 5 \\ 4\sqrt{2} & 3 & -3 \\ -3\sqrt{2} & 4 & -4 \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

3. (i) $\mathbf{P} = \begin{bmatrix} \dfrac{-1-i}{\sqrt{3}} & \dfrac{1+i}{\sqrt{6}} \\ \dfrac{1}{\sqrt{3}} & \dfrac{2}{\sqrt{6}} \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$

(ii) $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ -\dfrac{1-i}{\sqrt{6}} & 0 & \dfrac{1-i}{\sqrt{3}} \\ \dfrac{2}{\sqrt{6}} & 0 & \dfrac{1}{\sqrt{3}} \end{bmatrix}$; $\quad \mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{\mathrm{T}}\mathbf{AP} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

4. (i) $\lambda_1 \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \lambda_2 \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} = (1)\begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix} + (-1)\begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{bmatrix}$

(ii) $\lambda_1 \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \lambda_2 \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} = (5)\begin{bmatrix} \dfrac{2}{\sqrt{5}} \\ \dfrac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \end{bmatrix} + (-5)\begin{bmatrix} \dfrac{1}{\sqrt{5}} \\ -\dfrac{2}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \dfrac{1}{\sqrt{5}} & -\dfrac{2}{\sqrt{5}} \end{bmatrix}$

(iii) $\lambda_1 \mathbf{w}_1 \mathbf{w}_1^{\mathrm{T}} + \lambda_2 \mathbf{w}_2 \mathbf{w}_2^{\mathrm{T}} + \lambda_3 \mathbf{w}_3 \mathbf{w}_3^{\mathrm{T}}$

$$= (2)\begin{bmatrix} 0 \\ \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix} + (2)\begin{bmatrix} -\dfrac{1}{\sqrt{3}} \\ -\dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \end{bmatrix}\begin{bmatrix} -\dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{3}} \end{bmatrix} + (8)\begin{bmatrix} \dfrac{2}{\sqrt{6}} \\ -\dfrac{1}{\sqrt{6}} \\ \dfrac{1}{\sqrt{6}} \end{bmatrix}\begin{bmatrix} \dfrac{2}{\sqrt{6}} & -\dfrac{1}{\sqrt{6}} & \dfrac{1}{\sqrt{6}} \end{bmatrix}$$

Answers to Exercise Set 5

1. (i) $ax_1^2 + (b+c)\,x_1x_2 + dx_2^2$ (ii) $ax_1^2 + bx_2^2 + cx_3^2 + 2hx_1x_2 + 2gx_1x_3 + 5fx_2x_3$

(iii) $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + \lambda_4 x_4^2$

2. (i) $\begin{bmatrix} 2 & -\dfrac{7}{2} \\ -\dfrac{7}{2} & 5 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & -2 & 4 \\ -2 & 2 & \dfrac{5}{2} \\ 4 & \dfrac{5}{2} & -7 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{bmatrix}$

3. (i) Positive semi-definite (ii) Positive definite (iii) Indefinite

Gujarat Technological University

B.E. SEM-II (All Branches) Examination, June 2010

Solved Mathematics–II Question Paper

Time: 3 hours
Total Marks: 70

Q. 1(a) Attempt any two of the following:

(i) Solve the following system for x, y and z:

$$-\frac{1}{x}+\frac{3}{y}+\frac{4}{z}=30, \qquad \frac{3}{x}+\frac{2}{y}-\frac{1}{z}=9, \qquad \frac{2}{x}-\frac{1}{y}+\frac{2}{z}=10. \qquad \textbf{[6]}$$

Solution: The matrix form of the given system is

$$\begin{bmatrix} -1 & 3 & 4 \\ 3 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{x} \\ \frac{1}{y} \\ \frac{1}{z} \end{bmatrix}=\begin{bmatrix} 30 \\ 9 \\ 10 \end{bmatrix}$$

The augmented matrix of the given system is

$$\begin{bmatrix} -1 & 3 & 4 & : & 30 \\ 3 & 2 & -1 & : & 9 \\ 2 & -1 & 2 & : & 10 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 3 & 4 & : & 30 \\ 0 & 11 & 11 & : & 99 \\ 0 & 5 & 10 & : & 70 \end{bmatrix} \sim \begin{bmatrix} -1 & 3 & 4 & : & 30 \\ 0 & 1 & 1 & : & 9 \\ 0 & 1 & 2 & : & 14 \end{bmatrix}$$

$$\sim \begin{bmatrix} -1 & 3 & 4 & : & 30 \\ 0 & 1 & 1 & : & 9 \\ 0 & 0 & 1 & : & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 3 & 0 & : & 10 \\ 0 & 1 & 0 & : & 4 \\ 0 & 0 & 1 & : & 5 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & : & 2 \\ 0 & 1 & 0 & : & 4 \\ 0 & 0 & 1 & : & 5 \end{bmatrix}$$

Thus, $\frac{1}{x} = 2,\ \frac{1}{y} = 4,\ \frac{1}{z} = 5$

Hence $x = \frac{1}{2},\ y = \frac{1}{4},\ z = \frac{1}{5}$

(ii) Find $\mathbf{A}^{-1}$ using row operations if $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. **[3]**

Solution: For the given matrix:

$$[\mathbf{A} : \mathbf{I}] = \begin{bmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ -1 & 1 & 1 & : & 0 & 1 & 0 \\ 0 & 1 & 0 & : & 0 & 0 & 1 \end{bmatrix}$$

$$R_2 \to R_2 + R_1 \quad \sim \begin{bmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 2 & : & 1 & 1 & 0 \\ 0 & 1 & 0 & : & 0 & 0 & 1 \end{bmatrix}$$

$$R_3 \to R_3 - R_2 \quad \sim \begin{bmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 2 & : & 1 & 1 & 0 \\ 0 & 0 & -2 & : & -1 & -1 & 1 \end{bmatrix}$$

$$R_2 \to R_2 + R_3 \quad \sim \begin{bmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & 0 & 0 & 1 \\ 0 & 0 & -2 & : & -1 & -1 & 1 \end{bmatrix}$$

$$R_3 \to \left(-\frac{1}{2}\right) R_3 \quad \sim \begin{bmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ 0 & 1 & 0 & : & 0 & 0 & 1 \\ 0 & 0 & 1 & : & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$$R_1 \to R_1 - R_3 \quad \sim \begin{bmatrix} 1 & 0 & 0 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & : & 0 & 0 & 1 \\ 0 & 0 & 1 & : & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Hence the inverse of the given matrix $\mathbf{A}$ is $\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$.

(iii) Find the standard matrices for the reflection operator about the line $y = x$ on R^2 and the reflection operator about the yz-plane on R^3. **[3]**

Solution: The transformation $T: R^2 \to R^2$ that maps each vector into its symmetric image about the line $y = x$ is a "reflection" operator about the line $y = x$ on R^2, that is,

$$T(x, y) = (y, x).$$

For the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\} = \{[1, 0], [0, 1]\}$ of R^2,

$$T(\mathbf{e}_1) = [0, 1] = 0(\mathbf{e}_1) + 1(\mathbf{e}_2) \qquad T(\mathbf{e}_2) = [1, 0] = 1(\mathbf{e}_1) + 0(\mathbf{e}_2)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

So the required standard matrix for T is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B \ \ [T(\mathbf{e}_2)]_B]$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The reflection operator about the yz-plane on R^3 is defined by

$$T(x, y, z) = (-x, y, z)$$

For the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$ of R^3,

$$T(\mathbf{e}_1) = [-1, 0, 0] = (-1)(\mathbf{e}_1) + 0(\mathbf{e}_2) + 0(\mathbf{e}_3)$$
$$T(\mathbf{e}_2) = [0, 1, 0] = 0(\mathbf{e}_1) + 1(\mathbf{e}_2) + 0(\mathbf{e}_3)$$
$$T(\mathbf{e}_3) = [0, 0, 1] = 0(\mathbf{e}_1) + 0(\mathbf{e}_2) + 1(\mathbf{e}_3)$$

$$[T(\mathbf{e}_1)]_B = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_2)]_B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad [T(\mathbf{e}_3)]_B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore the associated matrix of the given reflection operator is

$$[T]_{B,B} = [[T(\mathbf{e}_1)]_B \ \ [T(\mathbf{e}_2)]_B \ \ [T(\mathbf{e}_3)]_B]$$

$$= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Q. 1(b) Show that there is no line containing the points (1, 1), (3, 5), (–1, 6) and (7, 2). **[3]**

Solution: Suppose $mx + c = y$ is a line passing through the given points.

Therefore all the four given points satisfy the line's equation, that is,

$$\begin{aligned} m + c &= 1 \\ 3m + c &= 5 \\ -m + c &= 6 \\ 7m + c &= 2 \end{aligned}$$

The augmented matrix for the given system is

$$\begin{bmatrix} 1 & 1 & : & 1 \\ 3 & 1 & : & 5 \\ -1 & 1 & : & 6 \\ 7 & 1 & : & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & : & 1 \\ 0 & -2 & : & 2 \\ 0 & 2 & : & 7 \\ 0 & -6 & : & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & : & 1 \\ 0 & -2 & : & 2 \\ 0 & 0 & : & 9 \\ 0 & 0 & : & -11 \end{bmatrix}$$

The last two rows of the above matrix suggest that the system is inconsistent. That is, it has no solution. Therefore, there is no line passing through the given points.

Q. 1(c)

(i) Find all vectors in R^3 of Euclidean norm 1 that are orthogonal to the vectors

$$\mathbf{u}_1 = [1, 1, 1] \text{ and } \mathbf{u}_2 = [1, 1, 0].$$ **[2]**

Solution: Suppose $\mathbf{v} = [v_1, v_2, v_3]$ is a vector with Euclidean norm 1 and it is orthogonal to the vectors $\mathbf{u}_1 = [1, 1, 1]$ and $\mathbf{u}_2 = [1, 1, 0]$. Therefore, we have the following equations for $\mathbf{v}$,

$$\begin{aligned} \|\mathbf{v}\| = 1 \quad &\Rightarrow \quad v_1^2 + v_2^2 + v_3^2 = 1 \\ \mathbf{u}_1 \cdot \mathbf{v} = 0 \quad &\Rightarrow \quad v_1 + v_2 + v_3 = 0 \\ \mathbf{u}_2 \cdot \mathbf{v} = 0 \quad &\Rightarrow \quad v_1 + v_2 = 0 \end{aligned}$$

Thus $v_1 = \pm\dfrac{1}{\sqrt{2}},\ v_2 = \mp\dfrac{1}{\sqrt{2}},\ v_3 = 0$

Hence $\left[\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, 0\right]$ and $\left[-\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right]$ are the vectors which have norm 1 and are orthogonal to the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$.

(ii) Find the rank of the matrix $\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -8 \end{bmatrix}$ in terms of determinants. **[2]**

Solution: For the given matrix $\mathbf{A}$,

$$\det \mathbf{A} = \begin{vmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -8 \end{vmatrix}$$

$$= 2\begin{vmatrix} 2 & -1 & 3 \\ 2 & -1 & 3 \\ -6 & 3 & -8 \end{vmatrix} = 0 \qquad \text{(since two rows of the determinant are equal)}$$

Thus, det **A** < 3. So we have to check the determinant of order 2.

$$\det\left(\begin{vmatrix} -2 & 6 \\ 3 & -8 \end{vmatrix}\right) = -2 \neq 0$$

Therefore, a determinant of order 2 obtained from the given matrix is nonzero and the determinant of order 3 is zero. Hence the rank of the given matrix is 2.

(iii) Is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ in row-echelon form or reduced row-echelon form? **[1]**

Solution: The given matrix $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is in the reduced row echelon form.

Q. 2(a)

(i) What conditions must b_1, b_2 and b_3 satisfy in order for

$$x_1 + 2x_2 + 3x_2 = b_1, \quad 2x_1 + 5x_2 + 3x_3 = b_2, \quad x_1 + 8x_3 = b_3$$

to be consistent? **[4]**

Solution: The matrix form of the given system is

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

The augmented matrix of the given system is

$$\sim \begin{bmatrix} 1 & 2 & 3 & : & b_1 \\ 2 & 5 & 3 & : & b_2 \\ 1 & 0 & 8 & : & b_3 \end{bmatrix}$$

$$\begin{matrix} R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1 \end{matrix} \sim \begin{bmatrix} 1 & 2 & 3 & : & b_1 \\ 0 & 1 & -3 & : & b_2 - 2b_1 \\ 0 & -2 & 5 & : & b_3 - b_1 \end{bmatrix}$$

$$R_3 \to R_3 + 2R_2 \sim \begin{bmatrix} 1 & 2 & 3 & : & b_1 \\ 0 & 1 & -3 & : & b_2 - 2b_1 \\ 0 & 0 & -1 & : & b_3 + 2b_2 - 5b_1 \end{bmatrix}$$

$$\begin{matrix} R_2 \to R_2 - 3R_3 \\ R_3 \to (-1)\, R_3 \end{matrix} \sim \begin{bmatrix} 1 & 2 & 3 & : & b_1 \\ 0 & 1 & 0 & : & -5b_2 + 13b_1 - 3b_3 \\ 0 & 0 & 1 & : & -b_3 - 2b_2 + 5b_1 \end{bmatrix}$$

$$R_1 \to R_1 - 2R_2 - 3R_3 \quad \sim \quad \begin{bmatrix} 1 & 0 & 0 & : & -40b_1 + 16b_2 + 9b_3 \\ 0 & 1 & 0 & : & -5b_2 + 13b_1 - 3b_3 \\ 0 & 0 & 1 & : & -b_3 - 2b_2 + 5b_1 \end{bmatrix}.$$

Therefore

$$x_1 = -40b_1 + 16b_2 + 9b_3$$
$$x_2 = -5b_2 + 13b_2 + 3b_3$$
$$x_3 = -b_3 - 2b_2 + 5b_1$$

Hence for all real values of b_1, b_2 and b_3, the given system is a consistent system.

(ii) Is $T: R^3 \to R^3$ defined by $T(x, y, z) = (x + 3y, y, z + 2x)$ linear? Is it one-to-one, onto or both? Justify. **[3]**

Solution: The given linear transformation is

$$T(x, y, z) = (x + 3y, y, z + 2x)$$

Linearity of *T*:

Suppose $\mathbf{u}_1 = [x_1, y_1, z_1]$ and $\mathbf{u}_2 = [x_2, y_2, z_2]$.

Hence $$T(\mathbf{u}_1 + \mathbf{u}_2) = T(x_1 + x_2, y_1 + y_2, z_1 + z_2)$$
$$= ((x_1 + x_2) + 3(y_1 + y_2), y_1 + y_2, (z_1 + z_2) + 2(x_1 + x_2))$$
$$= ((x_1 + 3y_1) + (x_2 + 3y_2), y_1 + y_2, (z_1 + 2x_1) + (z_2 + 2x_2))$$
$$= (x_1 + 3y_1, y_1, z_1 + 2x_1) + (x_2 + 3y_2, y_2, z_2 + 2x_2)$$

Therefore, $$T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)$$

Now, $$T(\alpha \boldsymbol{u}) = T(\alpha x, \alpha y, \alpha z)$$
$$= ((\alpha x) + 3(\alpha y), (\alpha y), (\alpha z) + 2(\alpha x))$$
$$= (\alpha(x + 3y), \alpha y, \alpha(z + 2x))$$
$$= \alpha(x + 3y, y, z + 2x)$$

Therefore, $$T(\alpha \mathbf{u}) = \alpha T(\mathbf{u})$$

Hence *T* is linear.

One-to-oneness of *T*:

Suppose $$T(\mathbf{u}_1) = T(\mathbf{u}_2)$$
$$(x_1 + 3y_1, y_1, z_1 + 2x_1) = (x_2 + 3y_2, y_2, z_2 + 2x_2)$$
$$x_1 + 3y_1 = x_2 + 3y_2, \quad y_1 = y_2, \quad z_1 + 2x_1 = z_2 + 2x_2$$
$$x_1 = x_2, \quad y_1 = y_2, \quad z_1 = z_2$$

Hence *T* is a one-to-one transformation.

Ontoness of *T*:

Let $\mathbf{v} = [x', y', z']$ be any vector of co-domain space R^3. To check the ontoness of *T*, we have to find some vector $\mathbf{u} = [x, y, z]$ in R^3such that

$$T(\mathbf{u}) = \mathbf{v}$$

or $$(x + 3y, y, z + 2x) = (x', y', z')$$

$$x + 3y = x', \qquad y = y', \qquad z + 2x = z'$$
$$x = x' - 3y', \qquad y = y', \qquad z = z' - 2x' + 6y'$$
$$u = (x' - 3y', y', z' - 2x' + 6y') \in R^3$$ (since x', y' and z' are real numbers; so x, y and z are also real.)

Thus for any vector **v** of co-domain space, we got one vector **u** in R^3 such that $T(\mathbf{u}) = \mathbf{v}$. Hence the given transformation is onto.

Q. 2(b)

(i) Show that the set $S = \{e^x, xe^x, x^2e^x\}$ in $C^2(-\infty, \infty)$ is linearly independent. **[2]**

Solution: To check the linearity of a set of functions, we need to find the Wronskian of it. Here

$$f_1(x) = e^x, \qquad f_2(x) = xe^x, \qquad f_3(x) = x^2e^x$$
$$f'_1(x) = e^x, \qquad f'_2(x) = (x+1)e^x, \qquad f'_3(x) = (x^2 + 2x)e^x$$
$$f''_1(x) = e^x, \qquad f''_2(x) = (x+2)e^x, \qquad f''_3(x) = (x^2 + 4x + 2)e^x$$

The Wronskian $W(x)$ of these functions is

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_2''(x) & f_2''(x) & f_3''(x) \end{vmatrix}$$

$$= \begin{vmatrix} e^x & xe^x & x^2e^x \\ e^x & (x+1)e^x & (x^2+2x)e^x \\ e^x & (x+2)e^x & (x^2+4x+2)e^x \end{vmatrix}$$

$$= e^{3x}\begin{vmatrix} 1 & x & x^2 \\ 1 & x+1 & x^2+2x \\ 1 & x+2 & x^2+4x+2 \end{vmatrix}$$

$$= e^{3x}\begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 2 & 4x+2 \end{vmatrix}$$

$$= e^{3x}(4x + 2 - 4x) = 2e^{3x} \neq 0$$

This function $W(x)$ does not have the value zero for any x in the interval $(-\infty, \infty)$. Hence f_1, f_2, and f_3 form a linearly independent set.

(ii) Check whether $V = R^2$ is a vector space with respect to the operations

$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1 - 2, u_2 + v_2 - 3)$

and $\alpha(u_1, u_2) = (\alpha u_1 + 2\alpha - 2, \alpha u_2 - 3\alpha + 3), \quad \alpha \in R.$ **[5]**

Solution: Here we need to check the following ten properties of the vector space for $V = R^2$. Let $\mathbf{u} = [u_1, u_2]$, $\mathbf{v} = [v_1, v_2]$ and $\mathbf{w} = [w_1, w_2]$ be vectors of R^2.

1. ***Closure property for addition*** By the definition of addition,

$[u_1, u_2] + [v_1, v_2] = [u_1 + v_1 - 2, u_2 + v_2 - 3]$

Since $u_i, v_i \in R$. So $u_i + v_i \in R$, for each $i = 1, 2$.

Therefore $\mathbf{u} + \mathbf{v} \in R^3$.

2. ***Commutative property for addition***

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= [u_1 + v_1 - 2, u_2 + v_2 - 3]\\ &= [v_1 + u_1 - 2, v_2 + u_2 - 3] \qquad (\text{since } u_i + v_i = v_i + u_i \text{ in } R)\\ &= [v_1, v_2] + [u_1, u_2]\\ &= \mathbf{v} + \mathbf{u}\end{aligned}$$

i.e. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

3. ***Associative property for addition***

Let $\mathbf{w} = [w_1, w_2]$ be any vector in R^2. Then,

$$\begin{aligned}\mathbf{u} + (\mathbf{v} + \mathbf{w}) &= [u_1, u_2] + [v_1 + w_1 - 2, v_2 + w_2 - 3]\\ &= [u_1 + (v_1 + w_1 - 2) - 2, u_2 + (v_2 + w_2 - 3) - 3]\\ &= [(u_1 + v_1 - 2) + w_1 - 2, (u_2 + v_2 - 3) + w_2 - 3]\\ &= [(u_1 + v_1 - 2, u_2 + v_2 - 3)] + [w_1, w_2]\\ &= (\mathbf{u} + \mathbf{v}) + \mathbf{w}\end{aligned}$$

That is, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

Hence the addition operation is associative.

4. ***Zero vector*** If $\mathbf{0} = [0_1, 0_2]$ is a zero vector of $V = R^2$, then

$$\begin{aligned}\mathbf{u} + \mathbf{0} &= \mathbf{u}\\ [u_1, u_2] + [0_1, 0_2] &= [u_1, u_2]\\ [u_1 + 0_1 - 2, u_2 + 0_2 - 3] &= [u_1, u_2]\\ u_1 + 0_1 - 2 = u_1,\ u_2 + 0_2 - 3 &= u_2\\ 0_1 = 2, \quad 0_2 &= 3\end{aligned}$$

Thus, $\mathbf{0} = [0_1, 0_2] = [2, 3]$ is a zero vector for R^2 with respect to the given addition operation.

5. ***Additive inverse*** Suppose $\mathbf{u}' = [u_1', u_2']$ is an additive inverse of $\mathbf{u} = [u_1, u_2]$ in R^2, then

$$\begin{aligned}&\mathbf{u} + \mathbf{u}' = \mathbf{0}\\ &[u_1, u_2] + [u_1', u_2'] = [0_1, 0_2]\\ &[u_1 + u_1' - 2, u_2 + u_2' - 3] = [2, 3]\\ &u_1 + u_1' - 2 = 2, u_2 + u_2' - 3 = 3\\ &u_1' = 4 - u_1, u_2' = 6 - u_2\end{aligned}$$

Thus $\mathbf{u}' = [u_1', u_2'] = [4 - u_1, 6 - u_2]$ is an additive inverse of $\mathbf{u} = [u_1, u_2]$ in R^2 with respect to the given addition operation.

6. ***Closure for scalar multiplication*** By the definition of scalar multiplication

$$\alpha(u_1, u_2) = (\alpha u_1 + 2\alpha - 2, \alpha u_2 - 3\alpha + 3), \quad \text{for some scalar } \alpha \in R.$$

Since $u_i \in R$, $i = 1, 2$ and α is any real scalar, so $(\alpha u_i + 2\alpha - 2) \in R$ for each i, $i = 1, 2$.

Therefore if $\mathbf{u} = [u_1, u_2]$ is a vector of R^2, then $\alpha\mathbf{u}$ is also a vector R^2.

7. ***Distributive law for addition***

$$\begin{aligned}\alpha(u + v) &= \alpha[u_1 + v_1 - 2, u_2 + v_2 - 3]\\ &= [\alpha(u_1 + v_1 - 2) + 2\alpha - 2, \alpha(u_2 + v_2 - 3) - 3\alpha + 3]\\ &= [\alpha u_1 + \alpha v_1 - 2, \alpha u_2 + \alpha v_2 - 6\alpha + 3]\\ \alpha\mathbf{u} + \alpha\mathbf{v} &= [(\alpha u_1 + 2\alpha - 2, \alpha u_2 - 3\alpha + 3)] + [(\alpha v_1 + 2\alpha - 2, \alpha v_2 - 3\alpha + 3)]\\ &= [(\alpha u_1 + 2\alpha - 2) + (\alpha v_1 + 2\alpha - 2) - 2, (\alpha u_2 - 3\alpha + 3) + (\alpha v_2 - 3\alpha + 3) - 3]\\ &= [\alpha u_1 + \alpha v_1 + 4\alpha - 6, \alpha u_2 + \alpha v_2 - 6\alpha + 3]\end{aligned}$$

Therefore, $\alpha(\mathbf{u} + \mathbf{v}) \neq \alpha\mathbf{u} + \alpha\mathbf{v}$.

Hence the given addition is not distributive under the given scalar multiplication in R^2.

8. ***Distributive law for scalar multiplication*** Let β be any real scalar.

$$\begin{aligned}(\alpha + \beta)\mathbf{u} &= (\alpha + \beta)[u_1, u_2]\\ &= [(\alpha + \beta)u_1 + 2(\alpha + \beta) - 2, (\alpha + \beta)u_2 - 3(\alpha + \beta) + 3)\\ &= [\alpha u_1 + \beta u_1 + 2\alpha + 2\beta - 2, \alpha u_2 + \beta u_2 - 3\alpha - 3\beta + 3)\\ \alpha\mathbf{u} + \beta\mathbf{u} &= [(\alpha u_1 + 2\alpha - 2, \alpha u_2 - 3\alpha + 3) + (\beta u_1 + 2\beta - 2, \beta u_2 - 3\beta + 3)]\\ &= [(\alpha u_1 + 2\alpha - 2) + (\beta u_1 + 2\beta - 2) - 2, (\alpha u_2 - 3\alpha + 3) + (\beta u_2 - 3\beta + 3) - 3)\\ &= [\alpha u_1 + \beta u_1 + 2\alpha + 2\beta - 4, \alpha u_2 + \beta u_2 - 3\alpha - 3\beta + 3]\end{aligned}$$

Hence $$(\alpha + \beta)\mathbf{u} \neq \alpha\mathbf{u} + \beta\mathbf{u}.$$

9. ***Associative law for scalar multiplication*** Let m be any real scalar.

$$\begin{aligned}\alpha(\beta\mathbf{u}) &= \alpha[\beta u_1 + 2\beta - 2, \beta u_2 - 3\beta + 3]\\ &= [\alpha(\beta u_1 + 2\beta - 2) + 2\alpha - 2, \alpha(\beta u_2 - 3\beta + 3) - 3\alpha + 3]\\ &= [\alpha\beta u_1 + 2\alpha\beta - 2\alpha + 2\alpha - 2, \alpha\beta u_2 - 3\alpha\beta + 3\alpha - 3\alpha + 3]\\ &= [\alpha\beta u_1 + 2\alpha\beta - 2, \alpha\beta u_2 - 3\alpha\beta + 3]\end{aligned}$$

Also, $(\alpha\beta)\mathbf{u} = [\alpha\beta u_1 + 2\alpha\beta - 2, \alpha\beta u_2 - 3\alpha\beta + 3)$

Thus, $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$

10. $$\begin{aligned}1\mathbf{u} &= [1 \cdot u_1 + 2 \cdot 1 - 2, 1 \cdot u_2 - 3 \cdot 1 + 3]\\ &= [u_1, u_2]\\ &= \mathbf{u}\end{aligned}$$

Thus, $1\mathbf{u} = \mathbf{u}$

Hence R^2 is not a vector space over R with respect to given addition and scalar multiplication because it does not satisfy the distributive laws 7 and 8.

$$R_3 \leftrightarrow R_4 \quad \sim \quad \begin{bmatrix} 1 & 0 & 0 & : & 0 \\ 0 & 1 & 0 & : & 0 \\ 0 & 0 & 1 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Therefore the vector equation has only the trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Hence the set S_1 is linearly independent.

Q. 3(b)

(i) Find a standard basis vector that can be added to the set $S = \{[-1, 2, 3], [1, -2, -2]\}$ to produce a basis of R^3. **[3]**

Solution: Let $\mathbf{v}_1 = [-1, 2, 3]$, $\mathbf{v}_2 = [1, -2, -2]$.
It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set. Also

$$[S] = \text{span}\,(S) = \{\alpha(-1, 2, 3) + \beta(1, -2, -2) \mid \alpha, \beta \in R)\}$$
$$= \{(-\alpha + \beta, 2\alpha - 2\beta, 3\alpha - 2\beta) \mid \alpha, \beta \in R\}$$

Consider the standard basis vector $\mathbf{e}_1 = [1, 0, 0]$. Suppose

$$\mathbf{e}_1 = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2$$

i.e. $$[1, 0, 0] = \alpha[-1, 2, 3] + \beta[1, -2, -2]$$

or $$1 = -\alpha + \beta,\ \ 0 = 2\alpha - 2\beta,\ \ 0 = 3\alpha - 2\beta$$

$\therefore$ $$\alpha = 0, \beta = 0$$

Thus $\mathbf{e}_1 = [1, 0, 0]$ cannot be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Therefore, the set $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1\}$ is a linearly independent set of three vectors of R^3. Hence B is a basis for R^3.

(ii) Determine whether $\mathbf{b}$ is in the column space of $\mathbf{A}$, and if so, express $\mathbf{b}$ as a linear combination of the column vectors of $\mathbf{A}$ if $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$. **[3]**

Solution: Let $\mathbf{C}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{C}_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$ and $\mathbf{C}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ be the column vectors of a matrix $\mathbf{A}$.

Suppose a vector $\mathbf{b}$ is a linear combination of $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$.

i.e. $$\mathbf{b} = \alpha_1 \mathbf{C}_1 + \alpha_2 \mathbf{C}_2 + \alpha_3 \mathbf{C}_3 \qquad \text{for some scalar } \alpha_1, \alpha_2, \alpha_3.$$

The augmented matrix of the above system is

$$\begin{bmatrix} 1 & -1 & 1 & : & 2 \\ 1 & 1 & -1 & : & 0 \\ -1 & -1 & 1 & : & 0 \end{bmatrix}$$

$$\sim \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 2 & -2 & -2 \\ 0 & -2 & 2 & 2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:c} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The corresponding equations are

$$\alpha_1 = 1$$
$$\alpha_2 - \alpha_3 = -1$$

Therefore the system has infinitely many solutions. So **b** is in the column space of **A**.

Q. 3(c)

(i) If **A** is an $m \times n$ matrix, what is the largest possible value for its rank. **[1]**

Solution: Since rank $\mathbf{A} \le \min(m, n)$. The largest value of rank $\mathbf{A} = \min(m, n)$.

(ii) Find the number of parameters in the general solution of $\mathbf{Ax} = 0$ if **A** is a 5×7 matrix of rank 3. **[1]**

Solution: The number of parameters in the general solution of $\mathbf{Ax} = \mathbf{0}$ is 3.

OR

Q. 3(a)

(i) Find the basis and dimension of

$$W = \{(a_1, a_2, a_3, a_4) \in R^4 \mid a_1 + a_2 = 0,\ a_2 + a_3 = 0, a_3 + a_4 = 0\}$$ **[3]**

Solution: Here $W = \{(a_1, a_2, a_3, a_4)\ R^4 \mid a_1 + a_2 = 0, a_2 + a_3 = 0, a_3 + a_4 = 0\}$

$$= \{(a_1, -a_1, a_1, -a_1) \mid a_1 \in R\}$$
$$= \{a_1(1, -1, 1, -1) \mid a_1 \in R\}$$

Thus the set S is generated by the one vector $\mathbf{v} = [1, -1, 1, -1]$. That is, all the vectors of S are multiples of **v**. Therefore, $B = \{\mathbf{v}\}$ is a basis of W. Hence the dimension of W = number of vectors in $B = 1$.

(ii) Find the basis for the subspace of P_3 spanned by the vectors: $1 + x, x^2, -2 + 2x^2, -3x$. **[3]**

Solution: Consider a matrix **A** with $\mathbf{p}_1 = 1 + x$, $\mathbf{p}_2 = x^2$, $\mathbf{p}_3 = -2 + 2x^2$ and $\mathbf{p}_4 = -3x$ as row vectors, that is,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 0 & 2 \\ 0 & -3 & 0 \end{bmatrix}$$

So the subspace of P_3 spanned by $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ and $\mathbf{p}_4$ is same as the row space of **A**. Therefore, a basis for the row space of **A** is a basis for **V**. Reducing this matrix to row echelon form, we have

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore, the nonzero row vectors

$$\mathbf{r}_1 = [1, 0, 0] = 1 \quad \mathbf{r}_2 = [0, 1, 0] = x \quad \mathbf{r}_3 = [0, 0, 1] = x^2$$

form a basis for the row space of **A** and so it is a basis for the subspace of R^n spanned by the vectors $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ and $\mathbf{p}_4$.

Q. 3(b)

(i) Reduce $S = \{[1, 1, 0], [0, 1, -1], [0, 4, -3], [0, 2, 0]\}$ to obtain a basis of R^3. **[3]**

Solution: Let $\mathbf{v}_1 = [1, 1, 0]$, $\mathbf{v}_2 = [0, 1, -1]$, $\mathbf{v}_3 = [0, 4, -3]$ and $\mathbf{v}_4 = [0, 2, 0]$ be the vectors of R^3. Consider the vector equation of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = 0$$

i.e.
$$\alpha_1[1, 1, 0] + \alpha_2[0, 1, -1] + \alpha_3[0, 4, -3] + \alpha_4[0, 2, 0] = \mathbf{0}$$

$$\begin{aligned} \alpha_1 \qquad\qquad\qquad &= 0 \\ \alpha_1 + \alpha_2 + 4\alpha_3 + 2\alpha_4 &= 0 \\ -\alpha_2 - 3\alpha_3 \qquad &= 0 \end{aligned}$$

That is,
$$\alpha_1 = 0,\ \alpha_2 = -3\alpha_3,\ \alpha_4 = -\alpha_3/2$$

$$-3\alpha_3[0, 1, -1] + \alpha_3[0, 4, -3] - \frac{1}{2}\alpha_3[0, 2, 0] = \mathbf{0}$$

$$[0, 4, -3] = 3[0, 1, -1] + \frac{1}{2}[0, 2, 0]$$

Therefore the vector $\mathbf{v}_3$ is dependent, so we can remove it from the set S. After removing the vector $\mathbf{v}_3$ from the set S, we get new a set $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\} = \{[1, 1, 0], [0, 1, -1], [0, 2, 0]\}$. Consider a matrix **A** whose columns are $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_4$, that is,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\therefore \qquad \det \mathbf{A} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & -1 & 0 \end{vmatrix} = 1[(0) - (-2)] = 2 \neq 0$$

Therefore the column vectors of **A** are linearly independent. So $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_4$ are independent vectors. That is, the set S' is a linearly independent set with three vectors of R^3. Hence it is basis for R^3.

(ii) Find a basis for the row space of **A** and column space of **A** if $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 0 \end{bmatrix}$. Also verify the dimension theorem for matrices. **[3]**

Solution: Let $\mathbf{M}_1 = \left\{\begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}, \mathbf{M}_2 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \mathbf{M}_3 = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix}, \text{and } \mathbf{M}_4 \begin{bmatrix} 0 & 0 \\ -1 & 2 \end{bmatrix}\right\}$ be the vectors of M_{22}. The vector equation for S is

$$\alpha_1\mathbf{M}_1 + \alpha_2\mathbf{M}_2 + \alpha_3\mathbf{M}_3 + \alpha_4\mathbf{M}_4 = \mathbf{0} \qquad \text{for some scalars } \alpha_1, \alpha_2, \alpha_3, \alpha_4.$$

i.e.
$$\alpha_1\begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix} + \alpha_2\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + \alpha_3\begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} + \alpha_4\begin{bmatrix} 0 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{aligned} \alpha_1 \qquad\qquad\qquad\qquad &= 0 \\ 2\alpha_1 - \alpha_2 + 2\alpha_3 \qquad\quad &= 0 \\ \alpha_1 - \alpha_2 + 3\alpha_3 - \alpha_4 &= 0 \\ -2\alpha_1 \qquad + \alpha_3 + 2\alpha_4 &= 0 \end{aligned}$$

The matrix of the coefficient of the above equation is

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 2 & 0 \\ 1 & -1 & 3 & -1 \\ -2 & 0 & 1 & 2 \end{bmatrix}$$

$$\det \mathbf{A} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 2 & 0 \\ 1 & -1 & 3 & -1 \\ -2 & 0 & 1 & 2 \end{vmatrix} = 1\begin{vmatrix} -1 & 2 & 0 \\ -1 & 3 & -1 \\ 0 & 1 & 2 \end{vmatrix} = -1(6+1) - 2(-2) = -3 \neq 0$$

Therefore S is a linearly independent set.

Let $\mathbf{M} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ be any vector of M_{22}.

Suppose $\mathbf{M} = c_1\mathbf{M}_1 + c_2\mathbf{M}_2 + c_3\mathbf{M}_3 + c_4\mathbf{M}_4$ for some scalars c_1, c_2, c_3, c_4.

i.e.
$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = c_1\begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix} + c_2\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} + c_4\begin{bmatrix} 0 & 0 \\ -1 & 2 \end{bmatrix}$$

$$\begin{aligned} c_1 \qquad\qquad\qquad\qquad &= a_1 \\ 2c_1 - c_2 + 2c_3 \qquad\quad &= a_2 \\ c_1 - c_2 + 3c_3 - c_4 &= a_3 \\ -2c_1 \qquad + c_3 + 2c_4 &= a_4 \end{aligned}$$

The augmented matrix of the system is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 2 & -1 & 2 & 0 & : & a_2 \\ 1 & -1 & 3 & -1 & : & a_3 \\ -2 & 0 & 1 & 2 & : & a_4 \end{bmatrix}$$

By using the Gauss-elimination method,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 0 & -1 & 2 & 0 & : & a_2 - 2a_1 \\ 0 & -1 & 3 & -1 & : & a_3 - a_1 \\ 0 & 0 & 1 & 2 & : & a_4 + 2a_1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 0 & -1 & 2 & 0 & : & a_2 - 2a_1 \\ 0 & 0 & 1 & -1 & : & a_3 - a_2 + a_1 \\ 0 & 0 & 1 & 2 & : & a_4 + 2a_1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 0 & -1 & 2 & 0 & : & a_2 - 2a_1 \\ 0 & 0 & 1 & -1 & : & a_3 - a_2 + a_1 \\ 0 & 0 & 0 & 3 & : & a_4 + a_1 + a_2 - a_3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 0 & -1 & 2 & 0 & : & a_2 - 2a_1 \\ 0 & 0 & 1 & 0 & : & \dfrac{2a_3 - 2a_2 + 4a_1 + a_4}{3} \\ 0 & 0 & 0 & 1 & : & \dfrac{a_4 + a_1 + a_2 - a_3}{3} \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & : & a_1 \\ 0 & 1 & 0 & 0 & : & \dfrac{-7a_2 + 14a_1 + 4a_3 + 2a_4}{3} \\ 0 & 0 & 1 & 0 & : & \dfrac{2a_3 - 2a_2 + 4a_1 + a_4}{3} \\ 0 & 0 & 0 & 1 & : & \dfrac{a_4 + a_1 + a_2 - a_3}{3} \end{bmatrix}$$

$$\therefore \quad c_1 = a_1;\ c_2 = \frac{-7a_2 + 14a_1 + 4a_3 + 2a_4}{3};\ c_3 = \frac{2a_3 - 2a_2 + 4a_1 + a_4}{3};\ c_4 = \frac{a_4 + a_1 + a_2 - a_3}{3}$$

Thus S spans M_{22}. Therefore S is a basis for M_{22}.

Q. 4(a)

(i) Compute $d(f, g)$ for $f = \cos 2\pi x$ and $g = \sin 2\pi x$ in $V = C[0, 1]$ with inner product

$$\langle f, g \rangle = \int_0^1 f(x)\, g(x)\, dx.$$

Solution: $d(f, g) = \|f - g\|$

$$= \sqrt{\langle f - g, f - g \rangle} = \left[\int_0^1 (f - g)(x)\,(f - g)(x)\, dx \right]^{1/2}$$

$$= \left[\int_0^1 [(f-g)(x)]^2\,dx\right]^{1/2} = \left[\int_0^1 (\cos 2\pi x - \sin 2\pi x)^2 dx\right]^{1/2}$$

$$= \left[\int_0^1 (1 - 2\cos 2\pi x \sin 2\pi x)\,dx\right]^{1/2} = \left[\int_0^1 (1 - \sin 4\pi x)\,dx\right]^{1/2}$$

$$= \left[\left(x + \frac{\cos 4\pi x}{4\pi}\right)_0^1\right]^{1/2} = \left[\left(1 + \frac{\cos 4\pi}{4\pi} - \frac{1}{4\pi}\right)\right]^{1/2}$$

$$= 1$$

(ii) Find a basis for the orthogonal complement of the subspace of R^3 spanned by the vectors $\mathbf{v}_1 = [1, -1, 3]$, $\mathbf{v}_2 = [5, -4, -4]$ and $\mathbf{v}_3 = [7, -6, 2]$. **[3]**

Solution: Let W be a subspace of R^3 spanned by the given vector $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$. The space W is the same as the row space of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$$

From the above theorem, the null space of matrix **A** is the orthogonal complement of the row space of **A**, that is, the null space of matrix **A** is the orthogonal complement of the space W. Therefore, now we will try to find a basis of the null space of matrix **A**. Since the null space **A** is a solution space of the equation $\mathbf{AX} = \mathbf{0}$, we have

$$\begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} x_1 - x_2 + 3x_3 &= 0 \\ 5x_1 - 4x_2 - 4x_3 &= 0 \\ 7x_1 - 6x_2 + 2x_3 &= 0 \end{aligned}$$

To find the nonzero solution of the above system of linear equations, we use the Gauss-elimination method. The augmented matrix for the above system is

$$\begin{bmatrix} 1 & -1 & 3 & : & 0 \\ 5 & -4 & -4 & : & 0 \\ 7 & -6 & 2 & : & 0 \end{bmatrix} \quad \text{by applying the row operations: } R_2 - 5R_1,\ R_3 - 7R_1$$

$$\begin{bmatrix} 1 & -1 & 3 & : & 0 \\ 0 & 1 & -19 & : & 0 \\ 0 & 1 & -19 & : & 0 \end{bmatrix} \quad \text{by applying the row operations: } R_3 - R_2,\ R_1 + R_2$$

$$\begin{bmatrix} 1 & 0 & -16 & : & 0 \\ 0 & 1 & -19 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Therefore, we get

$$x_1 - 16x_3 = 0 \text{ and } \quad x_2 - 19x_3 = 0$$

i.e.

$$x_1 = 16x_3 \text{ and } x_2 = 19x_3$$

$$[x_1,\ x_2,\ x_3] = [16x_3,\ 19x_3,\ x_3]$$

$$[x_1,\ x_2,\ x_3] = x_3[16,\ 19,\ 1]$$

Therefore, the null space $N_A = \{k\,[16, 19, 1) \mid k \in R\}$
Hence $\{[16, 19, 1]\}$ is a basis for the orthogonal complement of the vector space spanned by the given vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$.

Q. 4(b)

(i) Let $W = \text{span}\left\{\left[\frac{4}{5}, 0, -\frac{3}{5}\right], [0, 1, 0]\right\}$. Express $\mathbf{w} = [1, 2, 3]$ in the form of $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathrm{w}_1 \in W$ and $\mathbf{w}_2 \in W^{\perp}$. **[3]**

Solution: Let $\mathbf{v}_1 = \left[\frac{4}{5}, 0, -\frac{3}{5}\right]$ and $\mathbf{v}_2 = [0, 1, 0]$. It is easy to check that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for the subspace W of R^3. Here $\mathbf{w} = [1, 2, 3] \in R^3$. we know that

$$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 \quad \text{where } \mathbf{w}_1 = \text{Proj}_w\,\mathbf{w} \in W \text{ and } \mathbf{w}_2 = \mathbf{w} - \text{Proj}_w\mathbf{w} \in W^{\perp}$$

Since $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for W, we can write

$$\begin{aligned}
\mathbf{w}_1 &= \langle \mathbf{w}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{w}, \mathbf{v}_2\rangle \mathbf{v}_2 \\
&= \left[(1)\left(\frac{4}{5}\right) + (2)\,(0) + (3)\left(\frac{3}{5}\right)\right]\left[\frac{4}{5}, 0, \frac{3}{5}\right] + [(1)\,(0) + (2)\,(1) + (3)\,(0)]\,[0, 1, 0] \\
&= \frac{13}{5}\left[\frac{4}{5}, 0, \frac{3}{5}\right] + 2[0, 1, 0] \\
&= \left[\frac{52}{25}, 0, \frac{39}{25}\right] + [0, 2, 0] \\
&= \left[\frac{52}{25}, 2, \frac{39}{25}\right]
\end{aligned}$$

$$\mathbf{w}_2 = \mathbf{w} - \mathbf{w}_1 = [1, 2, 3] - \left[\frac{52}{25}, 2, \frac{39}{25}\right] = \left[-\frac{27}{25}, 0, \frac{36}{25}\right]$$

Therefore $\quad \mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2 = \left[\frac{52}{25}, 2, \frac{39}{25}\right] + \left[-\frac{27}{25}, 0, \frac{36}{25}\right] = [1, 2, 3]$

where $\quad \mathbf{w}_1 = \left[\frac{52}{25}, 2, \frac{39}{25}\right] \in W$ and $\mathbf{w}_2 = \left[-\frac{27}{25}, 0, \frac{36}{25}\right] \in W^{\perp}$.

Therefore $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an eigenvector of matrix **A** corresponding to the eigenvalue $\lambda_1 = 2$. Hence the eigenspace corresponding to the eigenvalue $\lambda_1 = 2$ is

$$E(2) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}; \qquad \dim(E(2)) = 1.$$

So the geometric multiplicity of the eigenvalue $\lambda_1 = 2$ is 1, while the algebraic multiplicity is 2. That is, the algebraic multiplicity is not equal to the geometric multiplicity. Hence the given matrix **A** is not diagonalizable.

Q. 4(c) Show that P_3 and M_{22} are isomorphic. **[3]**

Solution: Consider the function φ: $P_3 \to M_{22}$ defined as

$$\varphi(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{bmatrix} a_0 & a_1 \\ a_2 & a_3 \end{bmatrix}$$

Here we will show that the above mapping is isomorphic between P_3 and M_{22}.

(i) Here φ is well-defined and one-to-one.

Let $\mathbf{p} = p_0 + p_1x + p_2x^2 + p_3x^3$ and $\mathbf{q} = q_0 + q_1x + q_2x^2 + q_3x^3$ be two vectors of P_3. Suppose

$$\varphi(\mathbf{p}) = \varphi(\mathbf{q})$$

$$\Leftrightarrow \qquad \varphi(p_0 + p_1x + p_2x^2 + p_3x^3) = \varphi(q_0 + q_1x + q_2x^2 + q_3x^3)$$

$$\Leftrightarrow \qquad \begin{bmatrix} p_0 & p_1 \\ p_2 & p_3 \end{bmatrix} = \begin{bmatrix} q_0 & q_1 \\ q_2 & q_3 \end{bmatrix}$$

$$\Leftrightarrow \qquad p_0 = q_0,\ p_1 = q_1,\ p_2 = q_2,\ p_3 = q_3$$

$$\Leftrightarrow \qquad p = q$$

Hence φ is one-to-one map between P_3 and M_{22}.

(ii) Here φ is onto.

It is obvious that for any vector $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of M_{22}, we have one corresponding vector $a + bx + cx^2 + dx^3$ in P_3 such that $\varphi(a + bx + cx^2 + dx^3) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Thus, φ is onto map.

Therfore φ is one-to-one and onto mapping between P_3 and M_{22}. Hence P_3 and M_{22} are isomorphic.

$$\therefore \qquad x_1 = \frac{18}{7}, \; x_2 = \frac{13}{7}$$

Therefore the solutions of the normal equation are $x_1 = \frac{18}{7}, x_2 = \frac{13}{7}$, i.e. $[x_1, x_2] = \left[\frac{18}{7}, \frac{13}{7}\right]$. Hence the least squares solution of $\mathbf{Ax} = \mathbf{b}$ is $\hat{\mathbf{x}} = \left[\frac{18}{7}, \frac{13}{7}\right]$. Also, the orthogonal projection of **b** on the column space of **A** is

$$\text{proj}_{C(\mathbf{A})}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{18}{7} \\ \frac{13}{7} \end{bmatrix} = \begin{bmatrix} \frac{31}{7} \\ -\frac{5}{7} \\ \frac{8}{7} \end{bmatrix}$$

(ii) Find the transition matrix from basis $B = \{[1, 0], [0, 1]\}$ of R^2 to basis $B' = \{[1, 1], [2, 1]\}$ of R^2. **[3]**

Solution: Say $\mathbf{u}_1 = [1, 0]$, $\mathbf{u}_2 = [0, 1]$ and $\mathbf{u}_1' = [1, 1]$, $\mathbf{u}_2' = [2, 1]$

By the definition of transition matrix, **P** is given by the formula

$$\mathbf{P} = [[\mathbf{u}_1]_{B'} \, [\mathbf{u}_2]_{B'}]$$

Since $\mathbf{u}_1 = [1, 0]$, $\mathbf{u}_2 = [0, 1]$

$$\mathbf{u}_1 = \alpha_1 \mathbf{u}_1' + \alpha_2 \mathbf{u}_2'; \qquad \mathbf{u}_2 = \beta_1 \mathbf{u}_1' + \beta_2 \mathbf{u}_2';$$

$$[1, 0] = \alpha_1[1, 1] + \alpha_2[2, 1] \qquad [0, 1] = \beta_1[1, 1] + \beta_2[2, 1]$$

$$[1, 0] = [-1]\,[1, 1] + (1)\,[2, 1] \qquad [0, 1] = (2)\,[1, 1] + (-1)\,[2, 1]$$

$$[\mathbf{u}_1]_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad [\mathbf{u}_2]_{B'} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Thus the transition matrix $\mathbf{P}' = [[\mathbf{u}_1]_{B'} \, [\mathbf{u}_2]_{B'}] = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$.

Q. 4(c) For the matrix $\mathbf{A} = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix}$, show that the row vectors form an orthonormal set in C^2.

Also, find A^{-1}. **[3]**

Solution: The row vectors for the given matrix $\mathbf{A} = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix}$ are

Q. 5(b) Find a matrix **P** that diagonalizes $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ and hence find $\mathbf{A}^{10}$. Also, find the eigenvalues of $\mathbf{A}^2$. **[4]**

Solution: First we have to find the matrix **P** which diagonalize **A**

Step 1 Eigenvalues of **A**

The characteristic equation of matrix **A** is

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} \lambda - 1 & 0 \\ 1 & \lambda - 2 \end{bmatrix}\right) = 0$$

or $$(\lambda - 2)(\lambda - 1) = 0$$

$\therefore$ $$\lambda_1 = 1, \lambda_2 = 2$$

Step 2 Eigenvectors of **A**

***Case* (i)** For the eigenvalue $\lambda_1 = 1$

Suppose $\mathbf{u} = [x_1, x_2] \in R^2$ is an eigenvector corresponding to eigenvalue $\lambda = 1$.

Therefore,

$$(\lambda \mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

or $$(\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$$

or $$\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let $$x_1 = x_2 = t; \quad \text{for any } t \in R$$

Then $$\begin{aligned} \mathbf{u} &= [x_1, x_2] \\ &= [t, t] \\ &= t[1, 1] \end{aligned}$$

Thus, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = 1$.

***Case* (ii)** For the eigenvalue $\lambda_2 = 2$

Suppose $\mathbf{v} = [y_1, y_2] \in R^2$ is an eigenvector corresponding to eigenvalue $\lambda = 1$.

Therefore,

$$(\lambda \mathrm{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

$$(2\mathbf{I} - \mathbf{A})\mathbf{v} = \mathbf{0}$$

or $$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\therefore$ $$y_1 = 0, y_2 = t; \quad \text{for any } t \in R$$

Then
$$\mathbf{v} = [y_1, y_2]$$
$$= [0, t]$$
$$= t[0, 1]$$

Thus, $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 2$.

Step 3 Construct $\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Then, $\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

And the diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\ \mathbf{AP} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

Step 4 Moreover,
$$\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1}$$
$$\mathbf{A}^{10} = \mathbf{PD}^{10}\mathbf{P}^{-1}$$
$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1^{10} & 0 \\ 0 & 2^{10} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ -1023 & 1024 \end{bmatrix}$$

Q. 5(c) Let T: $R^4 \to R^3$ be a linear transformation given by $T(x_1, x_2, x_3, x_4) = (w_1, w_2, w_3)$ where $w_1 = 4x_1 + x_2 - 2x_3 - 3x_4$, $w_2 = 2x_1 + x_2 + x_3 - 4x_4$, $w_3 = 6x_1 - 9x_3 + 9x_4$. Find the bases for the range and kernel of T. **[4]**

Solution: Let $\mathbf{v} = [x_1, x_2, x_3, x_4] \in R^4$

Suppose
$$\mathbf{v} \in \ker(T)$$
$$T(\mathbf{v}) = \mathbf{0}$$
$$T(x_1, x_2, x_3, x_4) = (0, 0, 0, 0)$$
$$(4x_1 + x_2 - 2x_3 - 3x_4,\ 2x_1 + x_2 + x_3 - 4x_4,\ 6x_1 - 9x_3 + 9x_4) = (0, 0, 0)$$
$$4x_1 + x_2 - 2x_3 - 3x_4 = 0$$
$$2x_1 + x_2 + x_3 - 4x_4 = 0$$
$$6x_1 - 9x_3 + 9x_4 = 0$$

The augmented matrix of the above system is
$$\begin{bmatrix} 4 & 1 & -2 & -3 & : & 0 \\ 2 & 1 & 1 & -4 & : & 0 \\ 6 & 0 & -9 & 9 & : & 0 \end{bmatrix}$$

$$(R_1 \leftrightarrow R_2) \quad \sim \quad \begin{bmatrix} 2 & 1 & 1 & -4 & : & 0 \\ 4 & 1 & -2 & -3 & : & 0 \\ 6 & 0 & -9 & 9 & : & 0 \end{bmatrix}$$

$$(R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1) \quad \sim \quad \begin{bmatrix} 2 & 1 & 1 & -4 & : & 0 \\ 0 & -1 & -4 & 5 & : & 0 \\ 0 & -3 & -12 & 21 & : & 0 \end{bmatrix}$$

$$(R_3 \to R_3 - 3R_2) \quad \sim \quad \begin{bmatrix} 2 & 1 & 1 & -4 & : & 0 \\ 0 & -1 & -4 & 5 & : & 0 \\ 0 & 0 & 0 & 6 & : & 0 \end{bmatrix}$$

$$(R_3 \to R_3/6) \quad \sim \quad \begin{bmatrix} 2 & 1 & 1 & -4 & : & 0 \\ 0 & -1 & -4 & 5 & : & 0 \\ 0 & 0 & 0 & 1 & : & 0 \end{bmatrix}$$

$$(R_2 \to R_2 - 5R_3,\ R_1 \to R_1 + 4R_3) \quad \sim \quad \begin{bmatrix} 2 & 1 & 1 & 0 & : & 0 \\ 0 & -1 & -4 & 0 & : & 0 \\ 0 & 0 & 0 & 1 & : & 0 \end{bmatrix}$$

$$(R_1 \to R_1 + R_2) \quad \sim \quad \begin{bmatrix} 2 & 0 & -3 & 0 & : & 0 \\ 0 & -1 & -4 & 0 & : & 0 \\ 0 & 0 & 0 & 1 & : & 0 \end{bmatrix}$$

Thus, $\quad 2x_1 - 3x_3 = 0, \qquad -x_2 - 4x_3 = 0,\ x_4 = 0$

$\therefore \quad x_1 = \dfrac{3}{2}x_3; \qquad x_2 = -4x_3$

$$\ker(T) = \{(x_1, x_2, x_3, x_4) \mid T(x_1, x_2, x_3, x_4) = (0, 0, 0, 0)\}$$

$$= \left\{\left(\frac{3}{2}x_3, -4x_3, x_3, 0\right) \middle| x_3 \in R\right\}$$

$$= \left\{x_3\left(\frac{3}{2}, -4, 1, 0\right) \middle| x_3 \in R\right\}$$

Hence $\left\{\left[\dfrac{3}{2}, -4, 1, 0\right]\right\}$ is a basis for the kernel of the given linear transformation.

The range space of T is

$$R(T) = \{T(\mathbf{v}) \mid \mathbf{v} = [x_1, x_2, x_3, x_4] \in R^4\}$$

$$= \{(4x_1 + x_2 - 2x_3 - 3x_4,\ 2x_1 + x_2 + x_3 - 4x_4,\ 6x_1 - 9x_3 + 9x_4) \mid x_1, x_2, x_3, x_4 \in R\}$$

$$= \{2x_1[2, 1, 3] + x_2[1, 1, 0] + x_3[-2, 1, -9] + x_4[-3, -4, 9] \mid x_1, x_2, x_3, x_4 \in R\}$$

Thus $S = \{[2, 1, 3], [1, 1, 0], [-2, 1, -9], (-3, -4, 9)\}$ is the set of generators of range space $R(T)$ of T. But

$$[-2, 1, -9] = 4[1, 1, 0] + 0\,[-3, -4, 9] - 3[2, 1, 3]$$

That means S is a linearly dependent set of vectors. By eliminating the vector [–2, 1, 9] from S, a new set B = {[2, 1, 3], [1, 1, 0], [–3, –4, 9]} is formed which is a linearly independent set and of span $R(T)$ and hence it is the basis for the range space $R(T)$ of T.

OR

Q. 5(c) Let $T: R^2 \rightarrow R^3$ be a linear transformation given by $T(x_1, x_2) = (x_2, -5x_1 + 13x_2, -7x_1 + 16x_2)$. Find the matrix for the transformation T with respect to the basis $B=\{[3, 1]^T, [5, 2]^T\}$ for R^2 and $B' = \{[1, 0, -1]^T, [-1, 2, 2]^T, [0, 1, 2]\}$ for R^3. **[4]**

Solution: Here $B = \{[3, 1]^T, [5, 2]^T\}$ is a basis of R^2. To find the matrix of the given linear transformation, first we have to find the images of vectors of basis B under T and then find the co-ordinates of the images with respect to basis $B' = \{[1, 0, -1]^T, [-1, 2, 2]^T, [0, 1, 2]\}$ for co-domain space R^3.

$$T(3, 1) = (1, -2, -5) \qquad T(5, 2) = (2, 1, -3)$$

Since $[1, 0, -1]^T$, $[-1, 2, 2]^T$ and $[0, 1, 2]^T$ are the basis vectors of co-domain space R^3 of T, so we will try to find the co-ordinates of $T(3, 1)$, $T(5, 2)$ with respect to basis B'. Consider the equation,

$$T(3, 1) = \alpha_1(1, 0, -1) + \alpha_2(-1, 2, 2) + \alpha_3(0, 1, 2)$$
$$T(5, 2) = \beta_1(1, 0, -1) + \beta_2(-1, 2, 2) + \beta_3(0, 1, 2)$$
$$(1, -2, -5) = (\alpha_1 - \alpha_2, 2\alpha_2 + \alpha_3, -\alpha_1 + 2\alpha_2 + 2\alpha_3);$$
$$(2, 1, -3) = (\beta_1 - \beta_2, 2\beta_2 + \beta_3, -\beta_1 + 2\beta_2 + 2\beta_3)$$

Hence $\quad \alpha_1 = 1, \alpha_2 = 0, \alpha_3 = -2; \quad \beta_1 = 3, \beta_2 = 1, \beta_3 = -1;$

$$[T(3,1)^T]_{B'} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}; \qquad [T(5,2)^T]_{B'} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$$

Therefore the associated matrix of linear transformation T is

$$\mathbf{A} = [[T(3,1)]_{B'}\ [T(5,2)]_{B'}] = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ -2 & -1 \end{bmatrix}.$$

Q. 5(b)

(i) Let $T: R^2 \rightarrow R^2$ be defined by $T(x, y) = (x + y, x - y)$. Is T one-one? If so, find the formula for $T^{-1}(x, y)$. **[4]**

Solution: To check the one-to-one of T, we have to show that the images of distinct vectors of R^2 (domain) under T are same, and the vectors are also the same.

Let $\mathbf{u} = [x_1, y_1)$, $\mathbf{v} = (x_2, y_2)$ be two vectors in $\mathbf{R}^2$.

Suppose $\quad T(\mathbf{u}) = T(\mathbf{v})$

$$T(x_1, y_1) = T(x_2, y_2)$$
$$(x_1 + y_1, x_1 - y_1) = (x_2 + y_2, x_2 - y_2)$$

Thus, $\quad x_1 + y_1 = x_2 + y_2; \quad x_1 - y_1 = x_2 - y_2$

$$\therefore \qquad x_1 = x_2; \quad y_1 = y_2$$

Hence $$[x_1, y_1] = [x_2, y_2]$$

or $$\mathbf{u} = \mathbf{v}$$

Therefore T is one-to-one map. Hence its inverse map T^{-1} exists.

Let $\mathbf{w} = (w_1, w_2)$ be a vector of $T(R^2)$. We want to find $\mathbf{v} = [x_1, x_2]$, a vector in R^2 such that

$$T(\mathbf{v}) = \mathbf{w}$$

$$T(x_1, x_2) = (w_1, w_2)$$

$$(x_1 + x_2, x_1 - x_2) = (w_1, w_2)$$

$$x_1 + x_2 = w_1, \; x_1 - x_2 = w_2$$

$$\therefore \qquad x_1 = \frac{w_1 + w_2}{2}; \quad x_2 = \frac{w_1 - w_2}{2}$$

$$\mathbf{v} = [x_1, x_2] = \left[\frac{w_1 + w_2}{2}, \frac{w_1 - w_2}{2}\right]$$

Therefore, the inverse of T is defined by the formula

$$T^{-1}(\mathbf{w}) = \mathbf{v}$$

$$T^{-1}(\mathbf{w}) = \left(\frac{w_1 + w_2}{2}, \frac{w_1 - w_2}{2}\right)$$

(ii) Find the eigenvalues of $\mathbf{A} = \begin{bmatrix} -420 & \frac{1}{2} & 576 \\ 0 & 0 & 0.6 \\ 0 & 0 & \sqrt{3} \end{bmatrix}$. Is $\mathbf{A}$ invertible? **[1]**

Solution: The given matrix is upper triangular matrix, so the diagonal entries -420, 0 and $\sqrt{3}$ are the eigenvalues of $\mathbf{A}$. Moreover, 0 is one of the eigenvalues of $\mathbf{A}$, so it is not an invertible matrix.

Q. 5(c) Translate and rotate the co-ordinates axes, if necessary, to put the conic $9x^2 - 4xy + 6y^2 - 10x - 20y = 5$ in standard position. Find the equation of the conic in the final co-ordinate system. **[4]**

Solution: The matrix representation of the quadratic equation is

$$9x^2 - 4xy + 6y^2 - 10x - 20y - 5 = [x \;\; y]\begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [-10 \;\; -20]\begin{bmatrix} x \\ y \end{bmatrix} - 5 = 0$$

Take $\mathbf{A} = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$ and $\mathbf{k} = [-10 \;\; -20]$

To construct a matrix $\mathbf{P}$ that is orthogonally diagonalized, we will try to find the orthonormal bases of each eigenspace of $\mathbf{A}$. The characteristic equation of $\mathbf{A}$ is

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

$$\det\left(\lambda\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}-\begin{bmatrix}9 & -2\\-2 & 6\end{bmatrix}\right)=0$$

$$\det\left(\begin{bmatrix}\lambda-9 & 2\\2 & \lambda-6\end{bmatrix}\right)=0$$

or $$\lambda^2-15\lambda+50=0$$

or $$(\lambda-10)(\lambda-5)=0$$

$\therefore$ $$\lambda=5,\ \lambda=10$$

Thus $\lambda_1 = 5$ and $\lambda_2 = 10$ are the eigenvalues of **A**. It is easy to find that $\mathbf{w}_1=\begin{bmatrix}\frac{1}{\sqrt{5}}\\\frac{2}{\sqrt{5}}\end{bmatrix}$ and $\mathbf{w}_2=\begin{bmatrix}-\frac{2}{\sqrt{5}}\\\frac{1}{\sqrt{5}}\end{bmatrix}$ are the orthogonal bases for the eigenspaces corresponding to eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 10$ respectively.

So we can construct the matrix **P** such that det **P** = 1

$$\mathbf{P}=[\mathbf{w}_1\ \ \mathbf{w}_2]=\begin{bmatrix}\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}}\\\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{bmatrix}$$

Rotate the co-ordinate axes of the xy-co-ordinate system by substituting $\begin{bmatrix}x\\y\end{bmatrix}$ in

$$\mathbf{X}=\mathbf{PX}'$$

and $$\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}}\\\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix}$$

in the given quadratic equation. Hence, we get

$$\mathbf{X}^T\mathbf{AX}+\mathbf{kX}-5=0$$

or $$(\mathbf{PX}')^T\mathbf{A}(\mathbf{PX}')+\mathbf{k}(\mathbf{PX}')-5=0$$

or $$\mathbf{X}'^T\mathbf{DX}'+\mathbf{k}(\mathbf{PX}')-5=0$$

where $\mathbf{D}=\mathbf{P}^T\mathbf{AP}=\begin{bmatrix}5 & 0\\0 & 10\end{bmatrix}$

or $$[x'\ \ y']\begin{bmatrix}5 & 0\\0 & 10\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix}+[-10\ \ -20]\begin{bmatrix}\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}}\\\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}x'\\y'\end{bmatrix}-5=0$$

or
$$5x'^2 + 10y'^2 - \frac{50}{\sqrt{5}}x' - 5 = 0$$

or
$$[x' - \sqrt{5}]^2 + |\, 2y'^2 - 6 = 0 \qquad \text{(II)}$$

To put the conic in standard position, translate the co-ordinate by substituting

$$x'' = x' - \sqrt{5},\ y'' = y'$$

in the above equation (II). Hence, we get

$$x''^2 + 2y''^2 = 6$$

or
$$\frac{x''^2}{6} + \frac{y''^2}{3} = 1$$

which is the equation of ellipse.

Index